# Selective translational regulation of premature termination codon containing mutant mRNAs



**Won Kyu Kim**

**Deparment of Medical Science**

**The Graduate School, Yonsei University**

# Selective translational regulation of premature termination codon containing mutant mRNAs

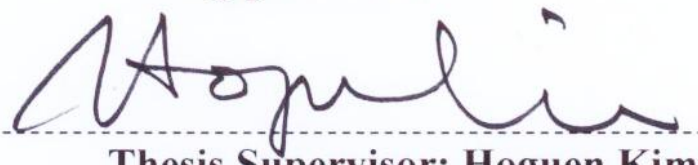
**Directed by Professor Hoguen Kim**

**The Doctoral Dissertation**
**submitted to the Department of Medical Science, the Graduate School of Yonsei University in partial fulfillment of the requirements for the degree of Doctor of Philosophy**

**Won Kyu Kim**

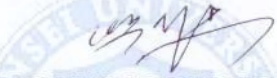
**June 2015**

**This certifies that the Doctoral Dissertation of Won Kyu Kim is approved.**

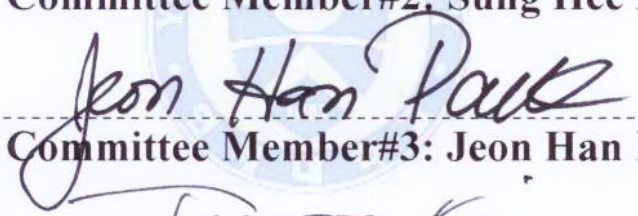


Thesis Supervisor: Hoguen Kim

Thesis Committee Member#1: Kyung-Sup Kim



Thesis Committee Member#2: Sung Hee Baek



Thesis Committee Member#3: Jeon Han Park

Thesis Committee Member#4: Jae-Woo Kim

**The Graduate School**

**Yonsei University**

**June 2015**

# ACKNOWLEDGEMENTS

지난 6 년간 공부하며 많은 잊지 못할 일들이 있었습니다. 항상 초심을 잃지 않고 성실하게 하고자 하는 마음을 유지하기 위해 스스로 다짐하고 노력했지만, 돌이켜 보면 과연 제가 충분히 열심히 했는지에 대해 걱정스러운 마음이 들기도 합니다. 그래서인지 지금도, 앞으로도 항상 부족하다고 생각하고 더욱 최선을 다 하고자 하는 마음이 간절합니다. 지난 6 년 동안 저의 능력 보다 많은 일을 할 수 있었던 이유는, 제가 힘들고 어려운 시기에 항상 저를 지지해 주고 도와주었던 감사한 분들 덕분인 것 같습니다. 먼저 지난 6 년간 항상 과학자로서 필요한, 또 세상을 살아가는 한 사람으로서 필요한 큰 가르침을 주시고 지도해 주신 김호근 교수님께 깊은 감사를 드리고 싶습니다. 아직 20 대의 어린 시절에 모르는 게 너무 많던 저를 제자로 받아주시고 지금의 모습으로 성장시켜 주셔서 항상 감사하고 죄송한 마음뿐 입니다. 그리고 학위 과정 동안 항상 좋은 조언과 독려해주신 김경섭 교수님, 백성희 교수님, 박전한 교수님, 그리고 김재우 교수님께 감사 드립니다. 개인적으로 오랜 시간 동안 저의 곁에서 항상 변함없는 모습으로 응원해주시고 지켜봐 주신 부모님과 너무나 사랑하는 동반자 정윤이, 그 동안 좋은 추억 많았던 랩 선배님들과 후배들 박민희, 윤성주, 조민희, 권유진, 박미선 그리고 전소연 모두 감사합니다. 제가 받은 사랑과 응원만큼 더욱 열심히 하고 싶습니다. 연구뿐만 아니라 저의 20 대부터 30 대 초반으로 넘어가는 인생을 함께한 랩에서 조금 더 잘 할 수 있었다는 아쉬움도 남지만, 한편으로는 앞으로 맞이할 새로운 연구와 삶이 더욱 기대됩니다. 앞으로 주어진 위치에서 겸손한 마음으로 항상 감사하며 성실하게 연구하는 과학자가 될 수 있기를 기도합니다.

김원규씀

# <TABLE OF CONTENTS>

**ABSTRACT.....1**

**I. INTRODUCTION.....3**

**II. MATERIALS AND METHODS.....6**

**1. Cell lines and tissue samples.....6**

**2. Identification of MSI and mutation analysis.....6**

**3. Semiquantitative RT-PCR and qRT-PCR.....6**

**4. Construction of expression plasmid vectors.....7**

**5. Transfection.....8**

**6. Western blotting and mutant MARKCKS protein-specific antibody generation.....8**

**7. Polysome assay.....9**

**8. Immunofluorescence microscopic examination.....9**

**9. Immunoprecipitation and ubiquitination assay.....10**

**10. Drug treatment and RNAi.....10**

**III. RESULTS.....11**

**1. Systematic search for the mutated genes containing PTC in the last exon.....11**

**2. Mutation profile of the genes containing coding mononucleotide repeats in the last exon in MSI-H cancer cell lines and tissues.....13**

**3. Expression of mutant mRNAs from genes containing PTCs in the last exon.....16**

**4. Analysis of endogenous truncated mutant proteins from NMD-irrelevant mutant mRNAs .....17**

**5. NMD-irrelevant mutant *MARCKS* mRNAs are associated with polysomes.....19**

**6. Endogenous NMD-irrelevant mutants of TTK, TCF7L2, and MARCKS proteins are generated, but mostly degraded by the proteasome system.....21**

**7. Overexpression of NMD-irrelevant mutants of MARCKS, TCF7L2, and TTK leads to heavy ubiquitination and localization around centrosomes.....24**

**8. Enhanced degradation of the truncated mutant MARCKS proteins containing neopeptides.....27**

**9. Confirmation of the existence of NMD-resistant PTC-containing mutant *β-globin* mRNAs from gBglo-P39 and gBglo-P66 constructs in the steady-state.....30**

**10. NMD-resistant PTC-containing mRNAs from gBglo-P39 and gBglo-P66 constructs are as stable as their PTC free counterparts but, mutant proteins are barely detected.....31**

**11. Overexpression of NMD-resistant mRNAs from the gBglo-P66 construct does not affect the generation of mutant proteins.....35**

**12. Mutant proteins from NMD-resistant PTC-containing mRNAs from the gBglo-P66 construct were mainly generated at the pioneer round of translation.....37**

**13. NMD-resistant PTC-containing mutant mRNAs from gBglo-P66 are not associated with polysomes.....38**

**14. Mutant proteins are actively translated from the rescued PTC-containing mRNAs from the gBglo-P66 construct when NMD is inhibited by UPF1 knock-down.....40**

**15. UPF1 plays key roles in the translational regulation of PTC-containing mRNAs.....42**

**16. NMD bypassed PTC-containing mRNAs from the CrPV-IRES-gBglo-P66 construct are translationally repressed by UPF1.....43**

**IV. DISCUSSION.....45**

**V. CONCLUSION.....50**

**REFERENCES.....52**

**ABSTRACT (IN KOREAN).....57**

**PUBLICATION LIST.....59**

# LIST OF FIGURES

**Figure 1. A pipeline for selecting candidate genes that potentially generate mutant proteins.....11**

**Figure 2. Mutation frequencies of the 35 genes containing coding mononucleotide repeats in MSI-H colon cancer cell lines and tumor tissues.....15**

**Figure 3. Semiquantitative RT-PCR and qRT-PCR analysis of the selected genes in 10 MSI-H and 2 MSS cell lines.....17**

**Figure 4. Measurement of protein expressions of endogenous mutant TTK, TCF7L2, and MARCKS.....19**

**Figure 5. Mutant mRNAs from the gMARCKS(-2) vector construct are associated with the heavy fractions of polysomes.....20**

**Figure 6. Endogenous mutant protein rescue assay using inhibitors of protein degradation pathways.....22**

**Figure 7. Endogenous mutant TTK, TCF7L2, and MARCKS are detected after proteasome inhibition.....24**

**Figure 8. Heavy ubiquitination of synthetic mutant TTK, TCF7L2, and MARCKS and their colocalization with centrosomes after proteasome inhibition.....25**

**Figure 9. Expression, degradation, and insolubility of wild-type and truncated mutant MARCKS containing or lacking neopeptides.....28**

**Figure 10. Precursor and mature RNA expressions from constructed human β-globin vectors.....31**

**Figure 11. PTC-containing mRNAs from gBglo-P39 and gBglo-P66 are stably expressed, but barely generate mutant proteins.....33**

**Figure 12. Transfection of increasing gBglo-P66 plasmids leads to the accumulation of NMD-resistant mRNAs, but the generation of mutant proteins is not changed.....36**

**Figure 13. Trace amount of mutant proteins from gBglo-P66 are mostly generated in the pioneer round of translation.....37**

**Figure 14. NMD-resistant mutant mRNAs from gBglo-P66 are translationally repressed.....39**

**Figure 15. Increased amount of mutant proteins from gBglo-P66 was detected when NMD was inhibited by UPF1 knock-down.....41**

**Figure 16. UPF1 plays key roles for the generation of a large amount of mutant proteins from the gBglo-P66 construct.....42**

**Figure 17. UPF1 represses translation of NMD bypassed PTC-containing mRNAs from CrPV-IRES-gBglo-P66.....44**

**Figure 18. Schematic model for the generation of mutant proteins from PTC-containing mRNAs.....51**

# LIST OF TABLES

**Table 1. Primers used for RT-PCR and qRT-PCR…….....7**

**Table 2. Primers used for the construction of expression vectors.....8**

**Table 3. List of genes containing coding mononucleotide repeats longer than 9 nt in the last exon…….....12**

**Table 4. Mutation profiles of 35 genes containing coding mononucleotide repeats in the last exon.....14**

**Table 5. Clinicopathologic features of 19 MSI-H colon cancer tissues and mutation profiles of three genes.....16**

**<ABSTRACT>**

**Selective translational regulation of premature termination codon containing mutant mRNAs**

**Won Kyu Kim**

***Department of Medical Science***

***The Graduate School, Yonsei University***

**(Directed by Professor Hoguen Kim)**

Abnormal mRNAs containing premature termination codons (PTCs) are normally degraded through nonsense-mediated mRNA decay (NMD) at the posttranscriptional level, thus preventing the production of potentially harmful truncated mutant proteins. However, NMD is known to be imperfect. The mRNAs containing PTCs in the last exon (NMD-irrelevant) are not recognized by NMD, so they are not degraded. About 10%-30% of NMD-competent PTC-containing mRNAs are resistant to NMD (NMD-resistant) and exist as stably as their PTC-free counterparts. Moreover, if NMD is inhibited by endogenous or exogenous ways, PTC-containing mRNAs to be degraded are expected to be rescued from NMD (rescued PTC-containing mRNAs). These NMD-irrelevant PTC-containing mRNAs, NMD-resistant PTC-containing mRNAs and rescued PTC-containing mRNAs from NMD are the potential sources for the generation of mutant proteins. Therefore, it has been urged to clarify whether mutant proteins are generated from the PTC-containing mRNAs, but detailed molecular mechanisms have been poorly understood.

In this study, to clarify whether mutant proteins from the NMD-irrelevant PTC-

containing mRNAs are produced, I identified 29 genes with frequent frameshift mutations in the last exon in colon cancers with high microsatellite instability (MSI-H) and selected three genes (*TTK*, *TCF7L2*, and *MARCKS*) for NMD-irrelevant PTC-containing mRNAs and protein expression analysis. The NMD-irrelevant PTC-containing mRNAs from these mutated genes were not degraded by NMD. However, only faint amounts of endogenous mutant TTK and TCF7L2 were detected, and I failed to detect endogenous mutant MARCKS. By polysome analysis, I demonstrated that NMD-irrelevant PTC-containing mRNAs are actively translated and inhibition of the proteasomal degradation facilitated the rescue of endogenous mutant TTK, TCF7L2, and MARCKS. By comparing the expression of neopeptide-containing or neopeptide-lacking truncated mutant proteins derived from genomic MARCKS constructs containing nonsense or frameshift mutations, I demonstrated that the enhanced degradation of these mutant proteins was driven by neopeptides. To demonstrate the generation of mutant proteins from NMD-resistant PTC-containing mRNAs and rescued PTC-containing mRNAs from NMD, human genomic β-globin expression constructs with nonsense mutations were used in order to avoid enhanced degradation of neopeptide-containing mutant proteins derived from frameshift mutations and I found that about 30 % of PTC-containing mRNAs are resistant to NMD in steady-state and these NMD-resistant PTC-containing mRNAs showed similar stability with their PTC-free counterparts. I demonstrated that the NMD-resistant PTC-containing mRNAs were translationally repressed and therefore, only trace amount of detected mutant proteins was generated from the pioneer round of translation step. Moreover, I found that PTC-containing mRNAs were significantly rescued when NMD was inhibited by down-regulating *hUPF1*, a key NMD factor and a large amount of mutant proteins from the rescued PTC-containing mRNAs was generated from the bulky translation step. I further demonstrated that down-regulation of UPF1 plays key roles to relieve the translational repression of PTC-containing mRNAs.

In conclusion, I found 1) NMD-irrelevant mutant mRNAs containing PTC in the

last exon are not degraded by NMD in cells and translated efficiently. 2) Truncated proteins containing neopeptides are rarely detected because of extensively degradation by the ubiquitin–proteasome system which is caused by neopeptides. 3) About 30% of PTC-containing mRNAs from NMD-competent β-globin expression constructs are resistant to NMD and exist as stably as their PTC free counterparts. 4) NMD-resistant PTC-containing mRNAs are translationally repressed and trace amount of mutant proteins from them is mainly generated from the pioneer round of translation step. 5) Mutant proteins from the rescued PTC-containing mRNAs from NMD by UPF1 down-regulation are mostly generated from bulky translation. 6) UPF1 plays key roles in the selective translational regulation of PTC-containing mRNAs depending on NMD status. These findings suggest that PTC-containing mRNAs are the pool of the sources for the mutant protein generation and are subject to differential translational and posttranslational regulations depending on NMD status.

Key words : truncated mutant protein, neopeptide, frameshift mutation, nonsense mutation, premature termination codon (PTC), nonsense-mediated mRNA decay (NMD), NMD-resistant, NMD-irrelevant, rescued PTC-containing mRNAs, *TTK*, *TCF7L2*, *MARCKS*, *β-globin*, *UPF1*

# Selective translational regulation of premature termination codon containing mutant mRNAs

**Won Kyu Kim**

***Department of Medical Science,***

***The Graduate School, Yonsei University***

**(Directed by Professor Hoguen Kim)**

## I. INTRODUCTION

Eukaryotic gene expression is controlled by a series of highly dynamic steps including pre-mRNA transcription, 5′ cap formation, splicing, polyadenylation, export, translation, and mRNA degradation. These steps are specifically integrated to augment the efficiency and fidelity of the entire process. Many proteins are bound to mRNA molecules throughout each step, and these mRNA-protein complex play roles in transport to the cytoplasm, subcellular localization, and translational control.[1-3] During the gene expression process, one of the most important quality control mechanisms is the quality control of transcribed mRNAs, nonsense mediated mRNA decay (NMD). The expressional inhibition of premature termination codon (PTC)-containing mRNAs, which are generated by nonsense and frameshift mutations, gene rearrangement, splicing, plays a predominant role in the mRNA quality control. If translated, PTC-containing mRNAs have the potential to

produce deleterious truncated proteins that could derange cellular functions through gain-of-function or dominant-negative activity.[4-7]

NMD is a quality control mechanism that rapidly detects and degrades PTC-containing mRNAs at the level of translation; thus, it acts as a surveillance mechanism to remove aberrantly processed and mutated mRNAs. Central to the NMD pathway, exon-junction complexes (EJCs) and UPF complexes play key roles.[8,9] When translation termination occurs at PTCs, the eukaryotic release factors eRF1 and eRF3 interact with UPF1 and SMG1 at the A-site of the terminating ribosome. This SMG1-UPF1-eRF complex associates with UPF2, which is recruited to EJC, and sequential phosphorylation of UPF1 by PI3-kinase-like SMG1 kinase recruits decay molecules such as SMG5, SMG6, and SMG7.[10] By these series of processes, most abnormal PTC-containing mRNAs are actively degraded, thus avoiding the potentially deleterious effects associated with the production of truncated proteins.[4,11]

Although NMD is a powerful RNA quality control mechanism, NMD is not a perfect mechanism. NMD is mediated through the recognition of PTC-containing mRNAs, which are recognized by their position relative to the last exon–exon junction. Mammalian transcripts that contain PTCs more than 50 to 55 nucleotides (nt) upstream of the last exon–exon junction are degraded by NMD, which ensures the degradation of most PTC-containing mRNAs. However, PTCs located within 50 to 55 nt or downstream of the last exon–exon junction are not recognized by NMD and can potentially lead to the generation of mutant proteins.[12,13]

A fundamental question is whether mutant proteins are generated from the PTC-containing mRNAs. To address this question is crucial because the generation of mutant proteins is directly linked to diseases such as genetic diseases and cancers.[14-16] A subset of colorectal tumors exhibits length alterations in several coding and noncoding microsatellites, a molecular phenotype termed high microsatellite instability (MSI-H).[17,18] These tumors are one of the ideal models to study whether

mutant proteins are generated from the PTC-containing mRNAs. The length alterations in microsatellites of the coding region [coding mononucleotide repeats (cMNRs)] result in nonsense or frameshift mutations in the affected genes and these mutations lead to the generation of PTC-containing mRNAs. Although many reports indicated that numerous genes are frequently mutated in their cMNRs in MSI-H cancers, few of these genes have been reported to express their mutant gene products.[14,19,20] Based on the idea that mRNAs containing PTCs in the last exon (NMD-irrelevant) are not recognized by NMD, it has been urged to reveal whether mutant proteins are generated from NMD-irrelevant mRNAs containing PTCs in the last exon .[14]

It is also important to determine if mutant proteins from mRNAs containing PTCs that are competent for NMD are generated in normal condition or condition of NMD inhibition because most reported PTC-containing mRNAs are competent for NMD. There are two possible ways for the generation of mutant proteins from NMD-competent mRNAs. First, mutant proteins can be generated from NMD-resistant PTC-containing mRNAs which are resistant to NMD by unknown mechanisms. It has been known that most of the PTC-containing mRNAs are instantly degraded by NMD, but 10%–30% PTC-containing mRNAs are resistant to NMD and exist as stably as their PTC-free counterparts.[21-24] If these 10%-30% mRNAs, so called NMD-resistant mRNAs, are translated, truncated mutant proteins can be generated. Second, a large amount of mutant proteins can be generated when NMD is inhibited because more PTC-containing mRNAs that are expected to be degraded will survive (rescued PTC-containing mRNAs). Based on these ideas, a study reported that mutant proteins can be used as tumor antigens for immunotherapeutic purposes. Inserting nucleotide sequences coding antigenic ovalbumin into β-globin reporter construct harboring PTC, it had been shown that tumors expressing mutant proteins from β-globin construct including ovalbumin coding region can be effectively removed by CD8+ T cells when NMD is inhibited.[25] However, the detailed molecular mechanisms in the generation of

mutant proteins from NMD-resistant and rescued PTC-containing mRNAs are poorly understood.

In this study, I addressed detailed molecular mechanisms for the generation of mutant proteins from PTC-containing mRNAs. To determine the generation of mutant proteins from NMD-irrelevant PTC-containing mRNAs, genes containing cMNRs in the last exon were obtained from public database and the mutation status was analyzed in these genes. Subsequently, I confirmed intact expressions of NMD-irrelevant PTC-containing mRNAs from the mutated genes containing cMNRs. It was demonstrated that mutant proteins were actively translated from genes containing mutations in cMNRs in the last exon but were rarely detected because these endogenous truncated proteins containing neopeptides were extensively degraded by the ubiquitin–proteasome system. It was also demonstrated that the enhanced degradation of the mutant proteins derived from NMD-irrelevant PTC-containing mRNAs was driven by neopeptdies. To determine the generation of mutant proteins from NMD-resistant and rescued PTC-containing mRNAs from NMD, I used human genomic β-globin expression vectors with nonsense mutations in order to avoid enhanced degradation of neopeptide-containing mutant proteins and found the stable existence of NMD-resistant PTC-containing mRNAs. It was demonstrated that the NMD-resistant PTC-containing mRNAs were translationally repressed and only trace amount of mutant proteins was generated from these NMD-resistant PTC-containing mRNAs in the pioneer round of translation step. It was further demonstrated that large amount of mutant proteins was generated from rescued PTC-containing mRNAs in the bulky translation step when NMD was inhibited by UPF1 down-regulation. Finally, I demonstrated that UPF1 plays key roles in the translational repression of NMD-competent (both NMD-resistant and rescued) PTC-containing mRNAs.

## II. MATERIALS AND METHODS

### 1. Cell lines and tissue samples

For the mutation analysis, 12 cell lines were used. DLD1, HCT116, HCT-8, LOVO, LS174T, RKO, SNUC2A, SNUC2B, SNUC4, and SNU407 cells are MSI-H colorectal carcinoma cell lines, whereas WiDr, HEK293, and HeLa cells are microsatellitestable (MSS) cell lines, as determined by previous studies.[26,27] Cells were grown in RPMI, minimum essential medium, and Dulbecco's modified Eagle medium supplemented with 10% FBS (Life Technologies, Waltham, Massachusetts, USA), 1% penicillin, and streptomycin at 37°C in 5% CO2. About tissue samples, 19 specimens confirmed as MSI-H colorectal carcinomas using BAT25, BAT26, D5S346, D17S25, and D2S123 were included in this study. The specimens were obtained from the archives of the Department of Pathology, Yonsei University (Seoul, Korea) and from the Liver Cancer Specimen Bank of the National Research Resource Bank Program of the Korea Science and Engineering Foundation of the Ministry of Science and Technology. Authorization for the use of the tissues for research was obtained from the Institutional Review Board of Yonsei University College of Medicine. Conventional pathologic parameters were examined without prior knowledge of the molecular data.

### 2. Identification of MSI and mutation analysis

Genomic DNA and cDNA preparation, analysis of MSI, and identification of target gene mutations were conducted using a PCR-based assay as described previously.[28]

### 3. Semiquantitative RT-PCR and qRT-PCR

The primers for semiquantitative reverse transcription (RT)-PCR and quantitative reverse transcription PCR (qRT-PCR) were designed using Primer 3 database (http://frodo.wi.mit.edu/primer3/). RNAs were isolated from cells using TRIzol (Invitrogen, Waltham, MA, USA). Reverse transcription was conducted using M-

MLV reverse transcriptase (Invitrogen). For RT-PCR, the reaction was conducted using Ampli-Taq Gold 360 DNA Polymerase (Applied Biosystems, Waltham, MA, USA). For qRT-PCR, the reaction was conducted using the ABI PRISM 7500 Sequence Detector (Applied Biosystems) and SYBR Premix Ex Taq II (TaKaRa, Seta, Shiga, Japan). The amount of target mRNA was normalized to that of GAPDH or EGFP mRNA [derived from the enhanced GFP (EGFP)-expressing control vector]. The sequences of the primers used are listed in Table 1.

**Table 1. Primers used for RT-PCR and qRT-PCR**

| Gene | Direction | Sequence |
|---|---|---|
| *GAPDH* | Forward | 5'- AAGGTGAAGGTCGGAGTCAAC |
| | Reverse | 5'- GGGGTCATTGATGGCAACAATA |
| *TTK* | Forward | 5'- TCATGCCCATTTGGAAGAGTC |
| | Reverse | 5'- CCACTTGGTTTAGATCCAGGC |
| *TCF7L2* | Forward | 5'- TGCGTTCGCTACATACAAGGT |
| | Reverse | 5'- AGGGGAGCCTAGCAGGTTC |
| *MARCKS* | Forward | 5'- AGCAACGAGACCCCGAAAA |
| | Reverse | 5'- CCTTCTCCAGCCTCCTTCTTGT |
| *EGFP* | Forward | 5'- ACGACGGCAACTACAAGACC |
| | Reverse | 5'- GTCCTCCTTGAAGTCGATGC |
| *ASTE1* | Forward | 5'- ATATGCCCCCGCTGAAATA |
| | Reverse | 5'- TTGGTGTGTGCAGTGGTTCT |
| *INO80E* | Reverse | 5'- CCTCCCACCCCCTAAGATG |
| | Reverse | 5'- ATCACCAGGTCATCGTCTCC |
| *CYHR1* | Forward | 5'- GGCCAACACTTATGGGATGT |
| | Reverse | 5'- GTAAGCCCAGCTGCCTACAG |
| *EBPL* | Forward | 5'- GTATGGCTGCTGGATGACCT |
| | Reverse | 5'- TCTAGCCATGACTGCCACAG |
| *β-globin* | Forward | 5'- GGCAACCCTAAGGTGAAGGC |
| | Reverse | 5'- GGTGAGCCAGGCCATCACTA |

**4. Construction of expression plasmid vectors**

For evaluating the generation of mutant proteins from NMD-irrelevant PTC-containing mRNAs, cDNA expression vectors for TTK [cTTK(WT)], TCF7L2 [cTCF4L2(WT)], and MARCKS [cMARCKS(WT)] were constructed by cloning the respective wild-type (WT) genes into pcDNA3.1 vectors containing a FLAG tag via amplification of their coding regions using the cDNAs derived from HeLa and WiDr cells. For the genomic DNA form of the MARCKS expression vector [gMARCKS(WT)], all exons and introns between the exons of *MARCKS* were cloned into pcDNA3.1 vectors containing a FLAG tag. To generate mutant protein expression vectors for TTK [cTTK(-2)], TCF7L2 [cTCF7L2(-1)], and MARCKS [cMARCKS(-2), gMARCKS(-2)], deletion mutagenesis was conducted. For the generation of mutant proteins from NMD-competent PTC-containing mRNAs, reporter vectors with the human originated *β-globin* gene were generated. So far, mouse-human β-globin reporter constructs have been used in many NMD studies that have focused on the effect of NMD on PTC-containing mRNA expression.[9,29,30] However, the mouse-human hybrid construct was not a proper model for studying the generation of mutant β-globin proteins from PTC-containing mRNAs because the artificial β-globin proteins from the hybrid constructs are not originated from human cells. These artificial proteins might be unstable by themselves or recognized as antigenic proteins in cells so they might subject to degradation, which makes it difficult to evaluate objective stability of β-globin mutant proteins.[31,32] To exclude unexpected influences on the generation of mutant proteins, human originated genomic *β-globin* gene was cloned into the 3xFLAG PCMV10 vector. Mutant β-globin expression vectors were generated by inserting nonsense mutations at specific positions (39$^{th}$, 66$^{th}$, 101$^{st}$ and 127$^{th}$) of β-globin amino acid. To construct CrPV-Bglo-WT and CrPV-Bglo-P66, a Nhe1 site was created upstream in the respective PCMV-B-globin plasmid by mutagenesis. CrPV-IRES sequences were PCR-amplified from pFR_CrPV_Xb (Add-gene plasmid 11509). The resulting PCR fragment was inserted into upstream of the gBglo-WT/P66 plasmids. The CMV10-

EGFP vector was used to confirm the transfection efficiency. The primers used for cloning are listed in Table 2.

**Table 2. Primers used for the construction of expression vectors**

| Gene | Direction | Sequence |
|---|---|---|
| pcDNA3.1 FLAG-TTK (cDNA form) | Forward | 5'- CCAGCGCAGCTTTCTGTAG |
| | Reverse | 5'- GATTTCCACAGGGATTCAAGA |
| pcDNA3.1 FLAG-TCF7L2 (cDNA form) | Forward | 5'- TGGCTTTTCTTCCTCCTTCA |
| | Reverse | 5'- AGCCACATGGCACAAAATTA |
| pcDNA3.1 FLAG-MARCKS (cDNA form) | Forward | 5'- TTCCCCTCTTGGATCTGTTG |
| | Reverse | 5'- AACAAACGGGGGAGAAAAGT |
| pcDNA3.1 FLAG-MARCKS (gDNA form) | Forward | 5'- TTCCCCTCTTGGATCTGTTG |
| | Reverse | 5'- ACCAGTACCTGGCACCACTC |
| PCMV10 3XFLAG-β-globin (gDNA form) | Forward | 5'- AGCAACCTCAAACAGACACC |
| | Reverse | 5'- GACCTCCCACATTCCCTTTT |
| PCMV10 3XFLAG-β-globin (cDNA form) | Forward | 5'- TGCAACCTCAAACAGACACCA |
| | Reverse | 5'- GCAAGAAAGCGAGCTTAGTGA |
| PCMV10 CrPV-IRES-3XFLAG-β-globin (gDNA form) | Forward | 5'- TGCTAGCACTAGTAAAGCAAAAATG |
| | Reverse | 5'- CATGGTGGCTAGCTTATCTTGA |

## 5. Transfection

All transfection experiments were carried out using Lipofectamine 2000 (Invitrogen).

## 6. Western blotting and mutant MARCKS protein-specific antibody generation

Whole lysates from cells were prepared using passive lysis buffer (Promega, Fitchburg, WI, United States). Membranes were incubated with primary antibodies against glyceraldehyde-3-phosphate dehydrogenase (GAPDH; Trevigen, MD, Gaithersburg, United States), FLAG (Sigma-Aldrich, St.Louis, MO, United States), TTK (Abnova, Taipei City, Taiwan), TCF7L2 (Cell Signaling Technology, Danvers,

MA, United States), MARCKS (Santa Cruz Biotechnology, Dallas, TX, United States), UPF1 (Cell signaling), EIF4A3 (Abnova), Y14 (Santa Cruz Biotechnology) and SMG1 (Santa Cruz Biotechnology) for 1 hour at room temperature. Antibodies against the neopeptide sequences of mutant MARCKS(-2) were generated in rabbits. All antibody generation procedures were conducted according to the manufacturer's manuals (Young In Frontier, Seoul, Korea).

**7. Polysome assay**

Twenty-four hours after plasmid transfection, HeLa cells were incubated with 100 μg/mL cycloheximide for 5 minuntes at room temperature and washed three times with ice-cold PBS. Cells were collected by scraping into PBS and incubated in lysis buffer [15 mM Tris-Cl, pH 7.4, 3 mM $MgCl_2$, 10 mM NaCl, 0.5% Triton X-100, 100 μg/mL cycloheximide, 1 mg/mL heparin, and 200 U RNasin (Intron, Gyeonggi-do, Korea)]. Where indicated, puromycin (100 μg/mL) was added to the cultures 2 hours prior to harvesting. Nuclei and debris were removed by centrifugation at 12,000 × *g* for 2 minutes. One milliliter of each sample was layered onto an 11 mL 10–50% sucrose gradient and centrifuged for 2 hours at 4°C using an SW40 rotor at 39,000 rpm. Eleven fractions were collected from the top of each gradient with concomitant measurements of absorbance at 254 nm using a fraction collection system. RNA was extracted from each fraction using TRIZOL reagent and analyzed by RT-PCR.

**8. Immunofluorescence microscopic examination**

Subcellular localization of mutant MARCKS was analyzed by immunofluorescence staining. The cells attached to glass coverslips were rinsed with PBS followed by fixation and permeabilization with ice-cold methanol for 10 minutes at -20 °C. Upon the removal of methanol, cells were again rinsed. Nonspecific sites were blocked with 2 % bovine serum albumin for 1 hour. After blocking, the medium was replaced with the respective primary antibodies, and cells were incubated

overnight. Cells were then washed and incubated for 1 hour with the appropriate fluorescently labeled secondary antibodies. For double-labeling experiments, cells were simultaneously incubated with the primary and secondary antibodies. Anti-FLAG M2-FITC (Sigma; F4049) and anti-γ-tubulin (Sigma; T5192) were used in this experiment. The Alexa Fluor® 594 dye conjugated secondary antibody (Life Technologies) was used for the γ-tubulin primary antibody. All images were obtained using an LSM700 confocal microscope (Carl Zeiss, Oberkochen, Germany).

**9. Immunoprecipitation and ubiquitination assay**

Immune complexes of wild-type and mutant MARCKS proteins were collected by gently rocking 1 mg of total proteins on an orbital shaker with prewashed anti-FLAG M2-agarose affinity gel (Sigma-Aldrich) at 4°C. The immune complexes bound to the affinity gel were washed and then boiled with a 100 mmol/L Tris–HCl–1% SDS solution to elute the complexes. Western blotting was conducted using FLAG and HA (Santa Cruz Biotechnology) antibodies. The relative density of each lane was quantified by ImageJ (NIH, Bethesda, Maryland, United States) software.

**10. Drug treatment and RNAi**

To inhibit proteasomal degradation, MG132 (25 uM) was treated for 6 hours before harvest. HeLa cells were treated with 50 ug/mL puromycin for 2 hours before harvest. Actinomycin D (50 ug/mL) was treated to block transcription for 4 to 8 hours before harvest. For the inhibition of bulky translation, 20 uM 4EGi inhibitors were treated for 6 hours before harvesting. HeLa cells were cultured in 60-mm dishes and were transiently transfected with 100 nM of UPF1, Y14, EIF4A3 or SMG1 siRNA (Bioneer) using Lipofectamine 2000 according to the manufacturer's protocol. Targeted nucleotides for UPF1 and Y14 were as follows: UPF1, 5'-(GAUGCAGUUCCGCUCCAUU) d(TT)-3', Y14, 5'- (CGC UCU GUU GAA

GGC UGG A) d(TT)-3'. Predesigned siRNAs for EIF4A3 [Sirna #(1040209) ] and SMG1 [Sirna#(1141625)] were purchased from Bioneer. Cells were then transfected with 1 μg of each specific plasmid as indicated in the text, 2 days after transfection with the siRNA. Cells were harvested 24 hours later and were used for protein purification and RNA extraction.

## III. RESULTS

### 1. Systematic search for the genes containing PTC in the last exon

I searched for human genes containing cMNRs longer than 9 nt in SelTarbase (http://www.seltarbase.org), a database providing comprehensive information about human mononucleotide microsatellite mutations and genes containing cMNRs. In total, 447 genes satisfying these criteria were obtained from SelTarbase. To confirm that these 447 genes include cMNRs longer than 9 nt, we analyzed the entire genes using bioinformatic tools such as Vector NTI (Invitrogen), the Human BLAST database (http://genome.ucsc.edu/), and the National Center for Biotechnology Information gene database. With this approach, 302 of 447 genes were confirmed to have cMNRs longer than 9 nt. On the basis of the NMD-irrelevant condition (mRNAs containing PTCs within 50–55 nt of the last exon junction), 302 genes were manually analyzed using Vector NTI software. When 1- or 2-bp deletions/insertions were detected in the cMNR region, genes that acquired abnormal stop codons distal to a site 50 to 55 nt from the last exon junction complex were selected. The number of finally selected genes was 66. Among them, 15 genes had previously been reported to have mutations in their cMNRs in MSI-H cancers (Fig. 1; http://www.sanger.ac.uk/genetics/CGP; refs. 19, 20). Most of the abnormal stop codons were PTCs; however, abnormal stop codons after normal stop codons (readthrough) were also found in some genes with cMNR mutations. Many of the 66 genes were related to biologically critical reactions, such as apoptosis,

cell-cycle regulation, cell proliferation, angiogenesis, and intracellular signaling (Table 3).

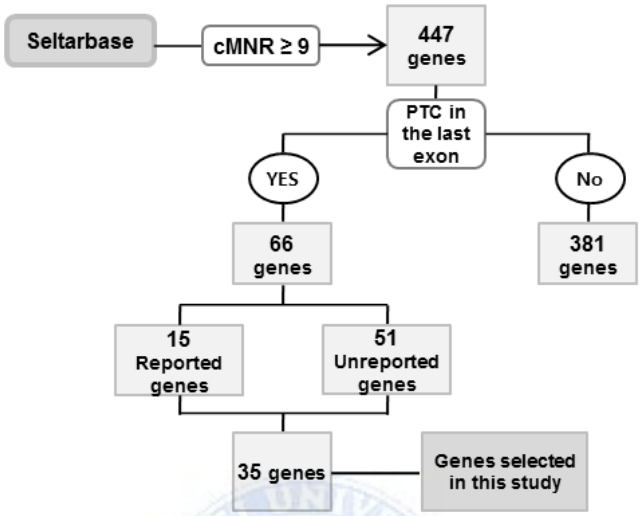



**Figure 1. A pipeline for selecting candidate genes that potentially generate truncated mutant proteins.**

**Table 3. List of genes containing coding mononucleotide repeats longer than 9 nt in the last exon**

| Gene | Repeat | Known function |
|---|---|---|
| AIM2* | A(10) | Tumor suppressor |
| ANKRD49 | A(9) | Unknown |
| ASH1L | A(9) | Histone methyltransferase |
| ASTE1* | A(11) | Possible role in EGF receptor signaling |
| BEND5 | A(9) | Unknown |
| C15orf48 | A(9) | Unknown |
| C4orf6* | T(10) | Unknown |
| CCDC43 | A(9) | Belongs to the CCDC43 family |
| CCT8L1 | A(12) | Possible molecular chaperone |
| CEL | C(9) | Catalyzes fat and vitamin absorption |
| CIR1(=UBE2V1) | A(9) | Unknown |
| CYHR1 | G(9) | Unknown |
| DENND1C | C(9) | Unknown |
| EBPL | T(9) | Unknown |
| ERCC5* | A(9) | Single-stranded structure-specific DNA endonuclease involved in DNA excision repair |

| | | |
|---|---|---|
| FAM111B | A(10) | Unknown |
| FBXL3 | T(9) | Probably recognizes and binds to some phosphorylated proteins and promotes their ubiquitination and degradation |
| FGFBP1 | A(9) | Acts as a carrier protein that release fibroblast-binding factors |
| FLT3LG* | C(9) | Stimulates the proliferation of early hematopoietic cells |
| GAFA1 | A(12) | Unknown |
| GBP3 | A(9) | Binds GTP, GDP and GMP |
| GINS1* | A(9) | The GINS complex plays an essential role in the initiation of DNA replication, and progression of DNA replication forks |
| GOT1L1 | A(9) | Unknown |
| HOXA11 | A(9) | Sequence-specific transcription factor which is part of a developmental regulatory system |
| INO80E | C(9) | Unknown |
| KCTD16 | A(9) | Unknown |
| KIAA1919 | T(9) | May function as a sodium-dependent glucose transporter |
| KIAA2018 | A(9) | Unknown |
| LOC100127950 | A(10) | Unknown |
| LOC100128175 | T(9) | Unknown |
| LOC100129345 | A(9) | Unknown |
| LOC100131089 | A(12) | Unknown |
| LOC643677 | A(9) | Unknown |
| MARCKS* | A(11) | The most prominent cellular substrate for protein kinase |
| MCHR2 | T(9) | Receptor for melanin-concentrating hormone, coupled to G proteins that activate phosphoinositide hydrolysis |
| OR52N5 | T(10) | Odorant receptor |
| OR6C76 | A(12) | Odorant receptor |
| OR7C1 | T(10) | Odorant receptor |
| OR7E24 | T(11) | Odorant receptor |
| PCDHB4 | A(9) | Potential calcium-dependent cell-adhesion protein |
| RBM43 | A(10) | Unknown |
| RG9MTD1 | A(10) | Functions in mitochondrial tRNA maturation |
| RGS12* | A(9) | Inhibits signal transduction by increasing the GTPase activity of G protein alpha subunits |
| RGS22 | A(9) | Inhibits signal transduction by increasing the GTPase activity of G protein alpha subunits |
| RTKN2 | A(9) | May play an important role in lymphopoiesis |
| RXFP2 | A(10) | Receptor for relaxin |
| SACS* | A(9) | May function in chaperone-mediated protein folding |
| SFRS12IP1* | A(10) | Possible splicing regulator involved in the control of cellular survival |
| SLAMF1 | A(9) | High-affinity self-ligand important in bidirectional T-cell to B-cell stimulation |
| SLC16A4* | T(9) | Proton-linked monocarboxylate transporter |
| SLC22A10 | A(9) | Unknown |
| SYCP1* | A(10) | Major component of the transverse filaments of synaptonemal complexe |

| | | |
|---|---|---|
| SYCP3* | A(9) | Component of the transverse filaments of synaptonemal complexes |
| TAS2R50 | T(9) | Receptor that may play a role in the perception of bitterness and is gustducin-linked |
| TBC1D23 | A(9) | Unknown |
| TCF7L2* | A(9) | Participates in the Wnt signaling pathway and modulates MYC expression by binding to its promoter in a sequence-specific manner |
| TMBIM4 | T(10) | Unknown |
| TMEM22 | A(9) | Unknown |
| TMEM60 | A(9) | Unknown |
| TMEM60 | A(9) | Unknown |
| TMEM97 | A(10) | Unknown |
| TRIM59 | A(10) | Unknown |
| TTK* | A(9) | Phosphorylates proteins on serine, threonine, and tyrosine and probably associated with cell proliferation and the mitotic checkpoint |
| VPS37B | C(9) | Component of the ESCRT-I complex, a regulator of vesicular trafficking process, maybe is involved in cell growth and differentiation. |
| ZNF292 | A(9) | Unknown |

***; Previously reported genes with mutations.**

## 2. Mutation profile of the genes containing cMNRs in the last exon in MSI-H cancer cell lines and tissues

I randomly selected 35 of 66 genes to analyze cMNR mutations by conducting an isotope PCR-based assay and sequencing. I used 10 MSI-H colon cancer cell lines for the mutation search. A MSS colon cancer cell line (WiDr) and the HeLa cell line were used as controls. We found frequent frameshift mutations in the cMNRs of 29 genes in the 10 MSI-H cell lines. The mutation profiles and status (homozygous vs. heterozygous) are summarized in Table 4. To validate the PCR-based mutation analysis, I conducted sequence analysis of several genes in the cell lines with homozygous mutations (Fig. 2A). All 10 MSI-H cancer cell lines had mutations in more than 4 of 35 genes examined, and 19 mutated genes were found in SNUC2A and SNUC2B cells (Fig. 2B). The mutation frequency of the 29 genes varied from 10% to 90% in the 10 MSI-H cell lines (Fig. 2C). I selected TTK, TCF7L2, and MARCKS and conducted PCR-based mutation analysis in 19 MSI-H colon cancer tissues. All 3 genes displayed frequent and varying mutation incidences ranging from approximately 35% to 60% (Fig. 2D). Information about the tissue samples used in this study is provided in Table 5.

**Table 4. Mutation profiles of human colon cancer cell lines in 35 genes containing coding mononucleotide repeats in the last exon**

| | MSI-H cancer cell line | | | | | | | | | | MSS cancer cell line | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Gene** | DLD1 | HCT116 | HCT8 | LOVO | LS174T | RKO | SNUC2A | SNUC2B | SNUC4 | SNU407 | WiDr | Hela |
| **Aim2** | -1/w* | -1/w | w | -1/-1 | -1/w | +1/-1 | -1/-2 | -1/-2 | w | w | w | w |
| **ASTE1** | +1/w | -1/w | +1/w | w | w | -2/-2 | -2/-2 | -2/-2 | -1/-1 | -2/-2 | w | w |
| **ASH1L** | w | w | w | w | w | w | -1/-1 | -1/-1 | -1/w | w | w | w |
| **ANKRD49** | w | w | -1/w | w | w | w | w | w | w | w | w | w |
| **BEND5** | w | w | w | w | -1/w | w | w | w | w | w | w | w |
| **CCDC43** | w | w | w | -1/w | -1/w | w | -1/w | -1/w | w | +1/w | w | w |
| **CCT8L1** | w | -1/-2 | w | -1/-2 | -1/-2 | -1/-2 | -2/-2 | -2/-2 | +1/-3 | -2/w | w | w |
| **CIR1** | w | w | w | w | w | w | w | w | w | w | w | w |
| **CYHR1** | w | w | w | w | -1/-1 | -1/w | -1/w | -1/w | +1/+2 | -1/w | w | w |
| **EBPL** | w | +2/w | w | -1/+1 | w | -1/w | w | -1/w | w | w | w | w |
| **ERCC5** | w | w | w | w | w | w | w | w | w | w | w | w |
| **FAM111B** | -1/w | -1/w | w | -1/-1 | -1/w | -1/-1 | -2/w | +1/-2 | -1/-1 | -1/-2 | w | w |
| **FBXL3** | w | w | w | w | w | w | w | w | -1/w | w | w | w |
| **FLT3LG** | w | w | w | +1/w | w | +1/w | +1/w | w | +1/w | w | w | w |
| **GAFA1** | w | -2/-2 | w | -2/-2 | -2/-2 | -2/-2 | -2/-3 | -2/-3 | -1/-1 | -1/-1 | w | w |
| **GBP3** | -1/w | -1/-1 | -1/w | w | -1/w | -1/w | w | w | -1/w | w | w | w |
| **GINS1** | w | w | w | w | w | -1/w | w | w | -1/w | w | w | w |
| **HOXA11** | w | w | w | w | w | w | w | w | w | w | w | w |
| **INO80E** | -1/w | w | -1/w | -1/-1 | -1/-1 | -1/-1 | -1/w | -1/w | -1/-1 | -1/w | w | w |
| **KCTD16** | w | -1/w | w | w | -1/w | w | -1/w | -1/w | w | -1/w | w | w |
| **KIAA2018** | w | w | w | w | w | w | w | w | w | -1/w | w | w |
| **LOC643677** | w | w | w | w | w | w | w | w | w | w | w | w |
| **LOC100127950** | w | w | w | -1/w | -1/w | w | +1/w | -1/w | -1/w | -1/w | w | w |
| **LOC100128175** | w | w | w | w | w | w | w | w | w | w | w | w |
| **LOC100131089** | w | -1/-1 | w | w | +1/+1 | -1/-1 | -1/-1 | -1/w | -1/w | w | w | w |
| **MED8** | w | w | w | -1/w | -2/W | w | -1/w | -1/w | -1/w | w | w | w |
| **MARCKS** | w | -1/w | w | -1/-1 | -1/-1 | -1/w | -2/w | -2/w | -1/-1 | -1/-1 | w | w |
| **RGS22** | w | w | w | -1/w | w | -1/w | w | w | w | w | w | w |
| **RXFP2** | w | w | w | w | w | w | w | w | w | w | w | w |
| **SYCP1** | -1/w | w | -1/w | +1/+1 | w | w | -1/-1 | -1/-1 | -1/-1 | -1/w | w | w |
| **SLAMF1** | w | -1/w | w | -1/w | w | w | w | w | -1/w | -1/w | w | w |
| **SFRS12IP1** | w | w | w | w | -1/-2 | w | +1/+1 | w | w | +1/+1 | w | w |
| **TRIM59** | w | w | w | -1/w | -1/w | -1/-1 | -1/w | -1/w | w | -2/-2 | w | w |
| **TCF7L2** | w | w | w | -1/w | -1/w | w | -1/w | -1/w | -1/-1 | -1/-1 | w | w |
| **TTK** | w | -1/w | w | -1/w | w | -2/w | -2/w | -2/w | +1/-2 | w | w | w |

The asterisk indicates mutation status of each gene: -1 denotes a 1-bp deletion in the cMNR, w denotes no mutation in the cMNR, and +1 denotes 1-bp insertion in the cMNR.

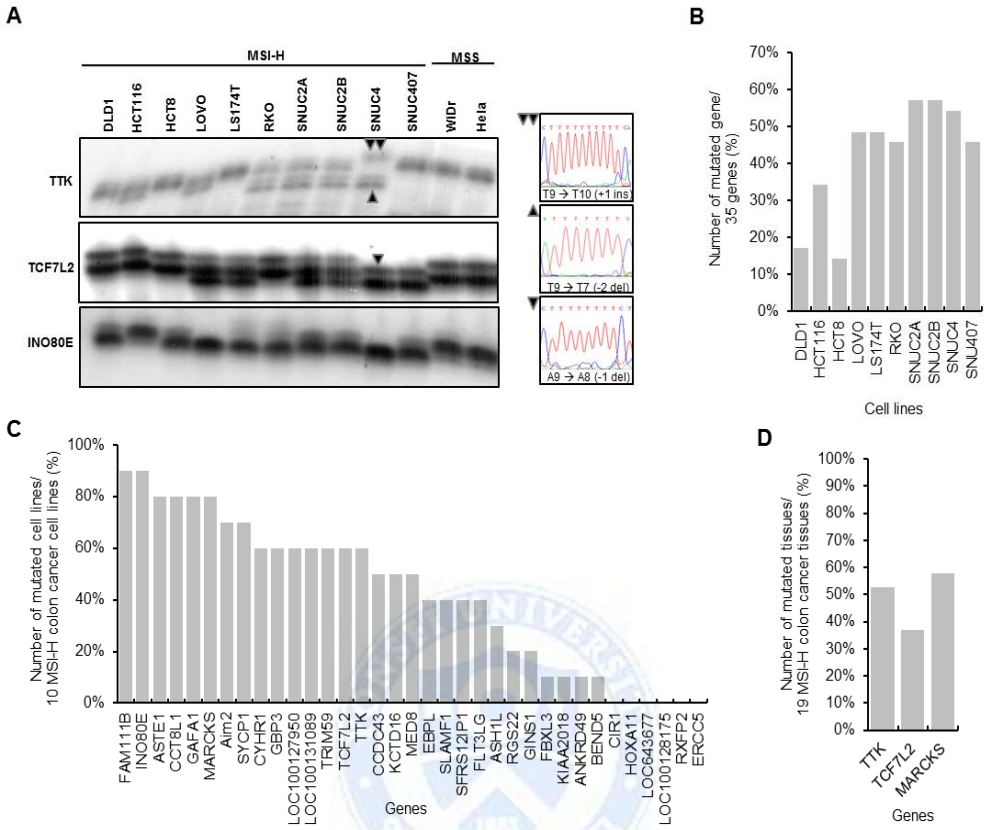



**Figure 2. Mutation frequencies of the 35 genes containing cMNRs in MSI-H colon cancer cell lines and tumor tissues. (**A) An example of the mutation analysis of TTK, TCF7L2, and MARCKS using a PCR-based assay and sequencing. Gel mobility shifts in the cells with 1-bp insertion (▼▼ ), 1-bp deletion ( ▲), and 2-bp deletions (▼ ) are evident (left). The DNA products displaying mobility shifts were confirmed as 1-bp insertions, 1-bp deletion, or 2-bp deletions in cMNR regions by sequence analysis. (B) The mutation frequencies of the 35 genes were analyzed in each cell line. The mean number of mutated genes in each cell line was 14.8 ± 5.43, and the mutation frequencies of 29 of these genes ranged from 10% to 60%. (C) Frequencies of mutated genes in 10 MSI-H cancer cell lines. (D) The mutation frequencies in the cMNRs of 3 genes (TTK, TCF7L2, and MARCKS) in 19 MSI-H colon cancer tissues.

**Table 5. Clinicopathologic features of 19 MSI-H colon cancer tissues and mutation profiles of three genes**

| Case Number | Sex | Age at diagnosis | Anatomic site | T | N | M | Stage | Tumor differentiation | Peritumoral lymphoid reaction | Mutation status | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | *TTK* | *TCF7L2* | *MARCKS* |
| 1 | F | 83 | Ascending | 3 | 0 | 0 | 2 | *MD | **1 | w | w | w |
| 2 | M | 75 | Sigmoid | 3 | 1 | 1 | 4 | PD | 1 | w | w | w |
| 3 | M | 71 | Ascending | 3 | 0 | 0 | 2 | MD | 2 | -1/w | w | -1/w |
| 4 | F | 38 | Ascending | 3 | 2 | 0 | 3 | MD | 2 | w | w | -1/w |
| 5 | M | 73 | Transverse | 3 | 0 | 0 | 2 | PD | 1 | w | w | w |
| 6 | F | 41 | Ascending | 3 | 0 | 0 | 2 | WD | 1 | w | w | w |
| 7 | F | 70 | Ascending | 3 | 0 | 0 | 2 | PD | 3 | w | w | -1/w |
| 8 | M | 71 | Ascending | 3 | 0 | 0 | 2 | MD | 1 | w | w | w |
| 9 | M | 60 | Ascending | 3 | 1 | 0 | 3 | MD | 2 | -1/w | w | -1/w |
| 10 | M | 47 | Ascending | 4 | 2 | 1 | 4 | MD | 2 | -1/w | -1/w | w |
| 11 | F | 72 | Ascending | 3 | 0 | 0 | 2 | MD | 2 | -1/w | -1/w | -1/w |
| 12 | M | 52 | Ascending | 3 | 0 | 0 | 2 | PD | 2 | -1/w | -1/w | -1 |
| 13 | M | 47 | Ascending | 2 | 0 | 0 | 1 | WD | 2 | -1/w | -1/w | w |
| 14 | M | 32 | Rectum | 2 | 0 | 0 | 1 | MD | 2 | -1/w | w | -1/w |
| 15 | F | 71 | Ascending | 3 | 0 | 0 | 2 | PD | 2 | w | -1/w | w |
| 16 | F | 58 | Sigmoid | 3 | 0 | 0 | 2 | PD | 2 | w | w | -1/w |
| 17 | F | 38 | Ascending | 1 | 0 | 0 | 1 | PD | 3 | -1/w | -1/w | -1 |
| 18 | M | 55 | Ascending | 3 | 0 | 0 | 2 | PD | 3 | -1/w | w | -1/w |
| 19 | F | 62 | Descending | 3 | 0 | 0 | 2 | MD | 3 | -1/w | -1/w | +1/w |

*; MD ; Moderate differentiation, WD ; Well differentiation, PD ; Poor differentiation

**; 1 ; Absent, 2 ; Mild, 3 ; Intense

### 3. Expression of mutant mRNAs from genes containing PTCs in the last exon

For the mutant protein expression analysis, I firstly selected 7 genes (TTK, TCF7L2, MARCKS, ASTE1, INO80E, CYHR1, and EBPL) displaying frameshift mutations in MSI-H cancer cell lines according to the availability of antibodies, presence of homozygous mutations, and cancer relevance. Before Western blot analysis, the mRNA expression of each gene was measured by qRT-PCR and semiquantitative

RT-PCR analysis. From these experiments, I sought to confirm that the mRNAs from mutated genes are not decayed by NMD or down-regulated by other genetic events, such as deletion or methylation. The mRNA expression of the selected genes was similarly quantified in the cell lines with homozygous mutations and compared with their levels in cells without mutations or those with heterozygous mutations in one allele (Fig. 3A). Importantly, SNUC4 cells, a MSI-H cell line with homozygous mutations in TTK (+1/-2), TCF7L2 (-1/-1), and MARCKS (-1/-1), exhibited relatively higher TTK, TCF7L2, and MARCKS mRNA levels than the other cell lines (Fig. 3B–D). These results indicate that mRNAs containing PTCs in the last exon are intact in cells irrespective of the mutation status of genes.

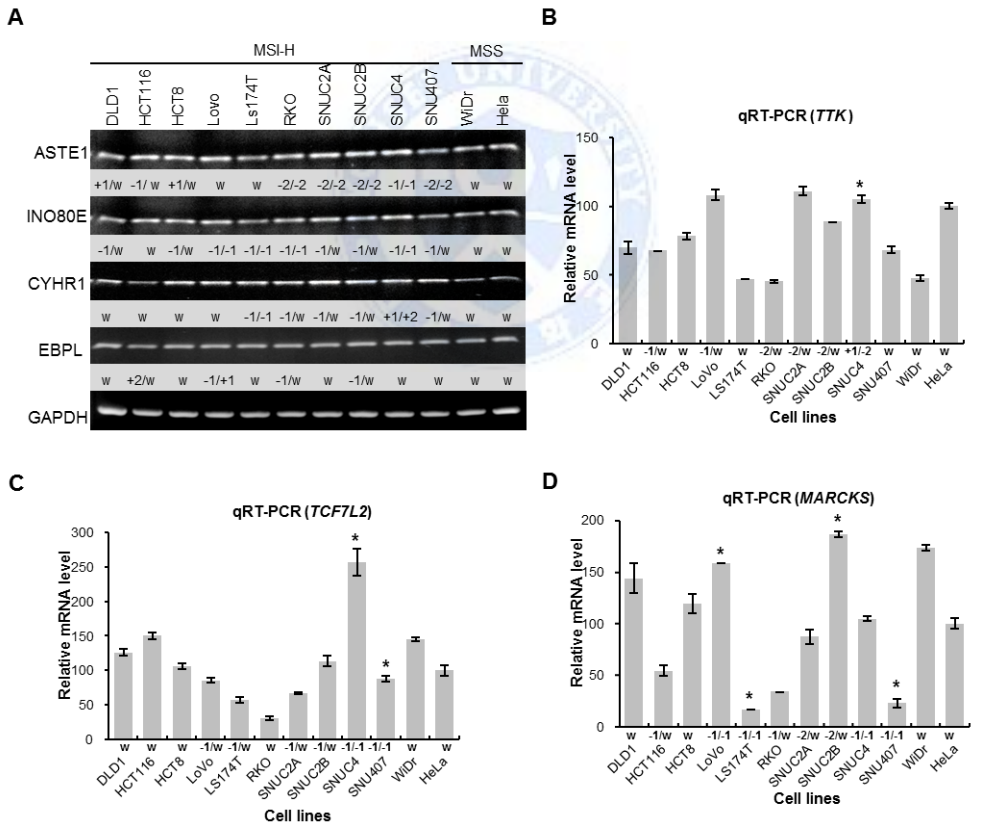



**Figure 3. Semiquantitative RT-PCR and qRT-PCR analysis of the selected genes in 10 MSI-H and 2 MSS cell lines.** TTK, TCF7L2, MARCKS, ASTE1, INO80E, CYHR1, and EBPL were selected for the mRNA expression study. (A) The mRNA expression levels of

ASTE1, INO80E, CYHR1, and EBPL were similar in these cell lines irrespective of their mutation status. (B–D) To precisely compare the mRNA expression levels of *TTK*, *TCF7L2*, and *MARCKS*, qRT-PCR was conducted using 10 MSI-H cell lines and 2 MSS cell lines. *, the mRNA expression level in the cell lines with homozygous mutation. -1, a 1-bp deletion in the cMNR; w, no mutation in the cMNR; +1, a 1-bp insertion in the cMNR; and -2, the 2-bp deletions in the cMNR.

**4. Analysis of endogenous truncated mutant proteins from mRNAs containing PTCs in the last exon**

I tested all the antibodies against the selected 7 genes by western blotting, and found that only antibodies against TTK, TCF7L2, and MARCKS showed excellent sensitivity and specificity for the experiments to follow. Then, I analyzed the expression of endogenous mutant TTK, TCF7L2, and MARCKS in 9 cell lines (7 MSI-H colon cancer cell lines, a MSS colon cancer cell line, and the HeLa cell line) by Western blotting. There was a minimal size difference between the mutant and normal proteins generated from TTK, whereas considerable size differences were observed in the normal and mutant proteins generated from TCF7L2 and MARCKS (Fig. 4A). I expected that if mutant proteins are expressed, then heterozygous mutation will cause quantitative and qualitative differences in the proteins and homozygous mutations will cause qualitative differences in the proteins because no wild-type proteins can be generated. To confirm this hypothesis, I quantified the normal TTK, TCF7L2, and MARCKS protein levels in each cell line based on their mutation status. Because of the heterozygous mRNA expression level in each cell line, I calculated the amount of normal protein by normalizing the protein level to the mRNA level, and then mean values were obtained depending on the mutation status. The result revealed that the normal protein level decreased according to the mutation status (wild-type > heterozygous mutation > homozygous mutations). For each mutant protein, cell lines with no mutations (4 cell lines) or heterozygous mutations (4 cell lines) in TTK exhibited a positive band at 90 kDa. The size difference between wild-type and mutant TTK was 0.3 kDa, and thus, normal and

mutant TTK cannot be distinguished by Western blotting in these cell lines. Therefore, SNUC4 cells represented the proper model for confirming whether mutant TTK is expressed because these cells have homozygous mutations in TTK. Although SNUC4 cells exhibited relatively higher TTK mRNA expression than the other cell lines used, only faint mutant TTK expression was detected. The size difference between wild type and mutant TCF7L2 was approximately 24 kDa. I detected faint bands only in cell lines with -1/-1 homozygous or -1/w heterozygous deletions (LS174T, SNUC4, and SNU407) compared with the patterns in cell lines with no mutations. The size difference between wild-type and mutant MARCKS was approximately 35 kDa. A complete loss of both normal and mutant MARCKS was observed in the 4 cell lines (LoVo, LS174T, SNUC4, and SNU407) with homozygous mutations (Fig. 4B). On the basis of these results, I suspected that either translation of the mutant mRNAs is repressed or mutant proteins are normally generated but extensively degraded through protein degradation pathways.

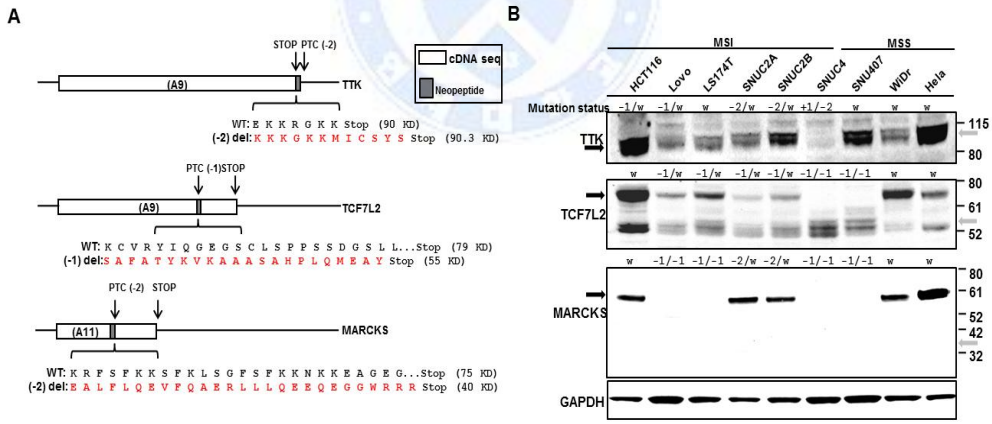



**Figure 4. Measurements of protein expressions of endogenous mutant TTK, TCF7L2, and MARCKS.** (A) Schematic diagram of *TTK*, *TCF7L2*, and *MARCKS*. The cDNA structure of each gene is presented using relative nucleotide numbers. The number and type of cMNR are represented as (A9) and (A11). (B) Western blotting was conducted using 7 MSI-H cancer cell lines and 2 MSS cancer cell lines with antibodies against TTK, TCF7L2,

and MARCKS. The expected sizes of wild type and mutant protein are marked by black and gray arrows, respectively. Trace amounts of TTK mutant were detected in SNUC4 cells with homozygous TTK mutations. Some faint bands consistent with the expected size of the TCF7L2 mutant were detected only in the cell lines with heterozygous (LS174T) and homozygous 1-bp deletions (SNUC4 and SNU407). MARCKS-mutant proteins were not detected at the expected size in the cell lines with heterozygous or homozygous deletion mutations (Lovo, LS174T, SNUC2A, SNUC2B, SNUC4, and SNU407).

## 5. mRNAs containing PTCs from mutant MARCKS are associated with polysomes

To evaluate the efficiency of the translation of mRNAs containing PTCs in the last exon, I generated the gMARCKS(WT) vector, a genomic DNA vector construct composed of 2 exons and 1 intron of MARCKS. I also generated the gMARCKS(-2) vector, a genomic DNA mutant MARCKS vector missing 2 adenine residues in the cMNR region (A11) of the last exon. I examined the translation of mutant mRNAs from the gMARCKS(-2) vector by analyzing the distribution of polysomes, using the gMARCKS(WT) vector as a control. Puromycin treatment was used to mimic a condition in which translation is repressed. The results revealed that wild-type *MARCKS* mRNA from HeLa cells transfected with the gMARCKS(WT) vector was mostly present in the polysome fractions (right shifted), similar to the *GAPDH* mRNA distribution, which indicated active translation (Fig. 5A). Mutant *MARCKS* mRNA-bearing PTCs from HeLa cells transfected with the gMARCKS(-2) vector exhibited a similar pattern as the wild-type *MARCKS* and endogenous *GAPDH* mRNAs (Fig. 5B). The normal translation of both wild-type and mutant *MARCKS* mRNAs was confirmed by the left shift in the banding pattern after puromycin treatment, indicating translational repression (Fig. 5C and 5D). On the basis of this polysome analysis, I concluded that both wild-type and mutant *MARCKS* mRNAs are actively translated.

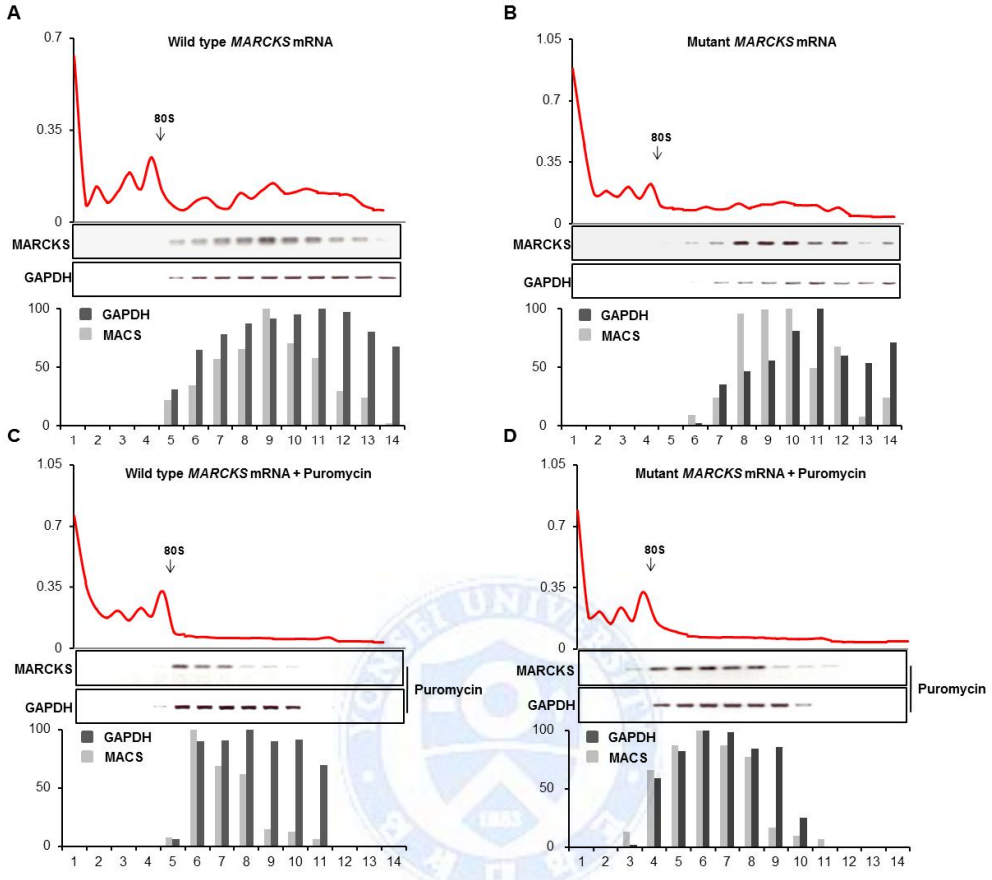



**Figure 5. Mutant mRNAs from the gMARCKS(-2) vector construct are associated with the heavy fractions of polysomes.** To confirm the translation efficiency of the mRNAs containing PTCs, vector constructs for the genomic DNA form of wild-type [gMARCKS(WT)] and mutant MARCKS [gMARCKS(-2)] were transfected into HeLa cells and polysome assay was performed. (A) The mRNAs from gMARCKS(WT) were mainly distributed in fractions 5 to 13, in which polysome peaks were observed. Most *GAPDH* mRNAs were also found in fractions 5 to 14, indicating that mRNAs from the gMARCKS(WT) construct are normally translated. (B) The mRNAs from gMARCKS(-2) were distributed in fractions 6 to 14, and *GAPDH* mRNAs were distributed in a similar pattern. (C) Cells were treated with puromycin to repress translation. The distributions of mRNAs from gMARCKS(WT) and *GAPDH* were clearly left-shifted, and the typical fluctuating polysome peaks were not detected, indicating the translational repression caused

by puromycin treatment. (D) After puromycin treatment, the distributions of mRNAs from gMARCKS(-2) and *GAPDH* were evidently left-shifted. The gentle and flat polysome peaks were also observed around heavy fractions. The intensities of all of the RT-PCR bands were measured, and they are shown as bars under the bands.

## 6. Endogenous mutant TTK, TCF7L2, and MARCKS are generated, but mostly degraded by the proteasome system

According to polysome analysis, I suspected that mutant proteins are generated but mostly degraded through proteolytic pathways. To determine which proteolytic pathways are responsible for degrading mutant proteins, I designed a rescue assay using several proteolytic inhibitors of 3 major protein degradation pathways: proteasome-, autophagy-, and lysosome-mediated degradation.[33,34] Bafilomycin A1 was used to block lysosomal degradation, and 3-methyladenine was used to inhibit autophagy mediated degradation. Lactacystin and MG132 were used to block the proteasomal pathway. For this experiment, I also generated mutant cDNA vectors of TTK [cTTK(-2)], TCF7L2 [cTCF7L2(-1)], and MARCKS [cMARCKS(-2)] to generate control mutant proteins (Fig. 6A). The rescue assays for mutant TTK and TCF7L2 were conducted using SNUC4 cells, and SNUC2B cells were used for mutant MARCKS. For the detection of mutant MARCKS, I generated a more sensitive and specific antibody against the neopeptide sequences in the C-terminal region of mutant MARCKS (Fig. 6B). This antibody more specifically and sensitively detected mutant MARCKS than the antibody recognizing the N-terminal region of MARCKS (Fig. 6C). The result showed that the levels of endogenous mutant TTK were significantly increased (approximately 2.5-fold) after proteasome inhibition, but no increment was observed when autophagy- or lysosome-mediated degradation was blocked. Expression of the TCF7L2 mutant was also greatly increased (4-fold) only after proteasome inhibition in SNUC4 cells. I conducted a mutant MARCKS rescue assay using the generated antibody. Interestingly, the expression of mutant MARCKS was increased by more than 100-fold when

proteasome degradation was blocked in SNUC2B cells (Fig. 6D).

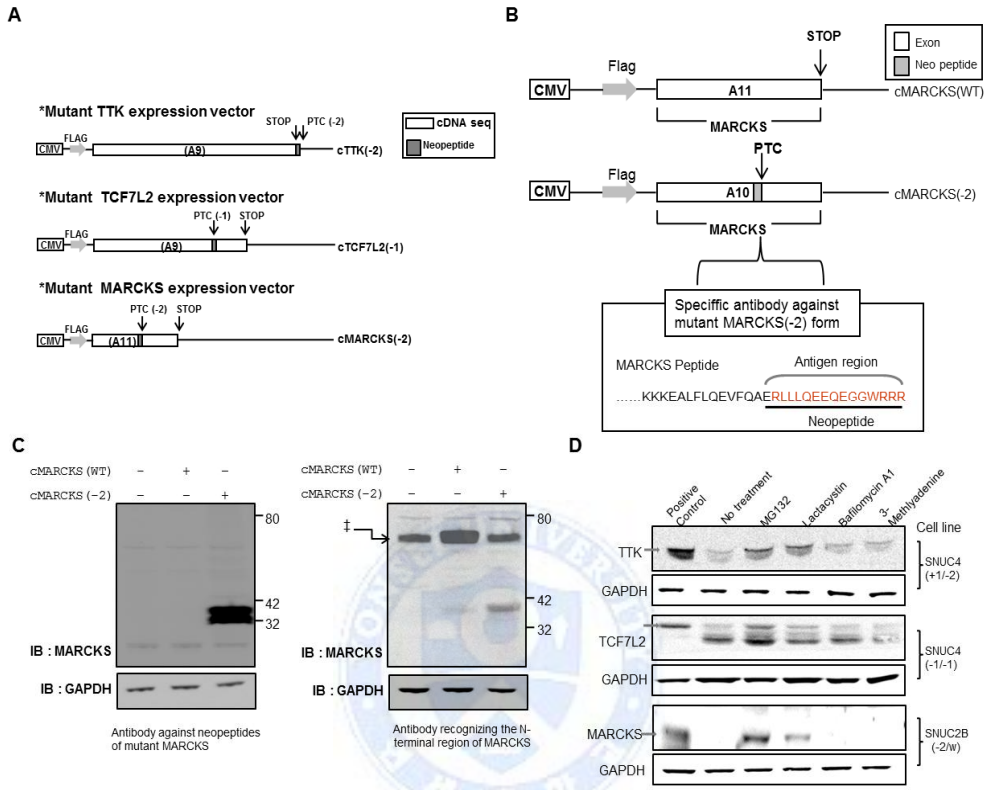



**Figure 6. Endogenous mutant protein rescue assay using inhibitors of protein degradation pathways.** (A) Schematic diagram of the wild-type and mutant TTK, TCF7L2, and MARCKS. (B) Schematic diagram of vector constructs used for peptide sources to generate antibodies. The antibody was generated against the antigen region (neopeptide sequences) in red. (C) To confirm the specificity and sensitivity of the antibodies, lysates from HeLa cells transfected with cMARCKS(WT) and cMARCKS(−2) were prepared and analyzed by Western blotting using a commercial antibody and the generated antibody. The MARCKS mutant-specific antibody exhibited very high sensitivity and specificity compared to those of the commercial antibody recognizing N-terminal region of MARCKS. (D) A mutant protein rescue assay using inhibitors of protein degradation pathways. Bafilomycin A1 was used to block lysosomal degradation, and 3-methyladenine was used to inhibit autophagy-mediated degradation. Lactacystin and MG132 were used to block the proteasomal pathway. The positive controls for each mutant were derived from HeLa cells

transfected with the respective mutant cDNA constructs. Gray arrows indicate the mutant protein sizes. Endogenous mutant TTK, TCF7L2, and MARCKS were increased after proteasome inhibition. -1, 1-bp deletion in the cMNR; -2, 2-bp deletion in the cMNR. ‡ denotes endogenous wild-type MARCKS protein.

I clarified these facts by using RNA interference (RNAi) against *TTK*, *TCF7L2*, and *MARCKS*. In SNUC4 cells treated with siRNA against *TTK*, the proteasome inhibition-induced increase in mutant TTK expression was reduced to approximately 45% of that in MG132-treated SNUC4 cells, confirming that these bands represent TTK. To confirm whether MG132 treatment specifically rescues mutant TTK, I conducted the same experiment using HeLa cells, which only express wild-type TTK. The results revealed that MG132 treatment specifically affects mutant TTK protein (Fig. 7A). As observed for mutant TTK, I showed that the proteasome inhibition–induced increases in mutant MARCKS and TCF7L2 expression were reduced by approximately 50% upon siRNA treatment. However, no changes were observed in HeLa cells after proteasome inhibition or siRNA treatment (Fig. 7B and 7C). I examined the stability of endogenous mutant TTK in a time course experiment using SNUC4 (+1/-2) and HeLa cells (wild-type). The expression of mutant TTK was dramatically decreased in the absence of MG132, which confirms that endogenous mutant TTK is rapidly degraded by the proteasome system (Fig. 7D). The expression of wild-type TTK was stable during the experiment (Fig. 7E).

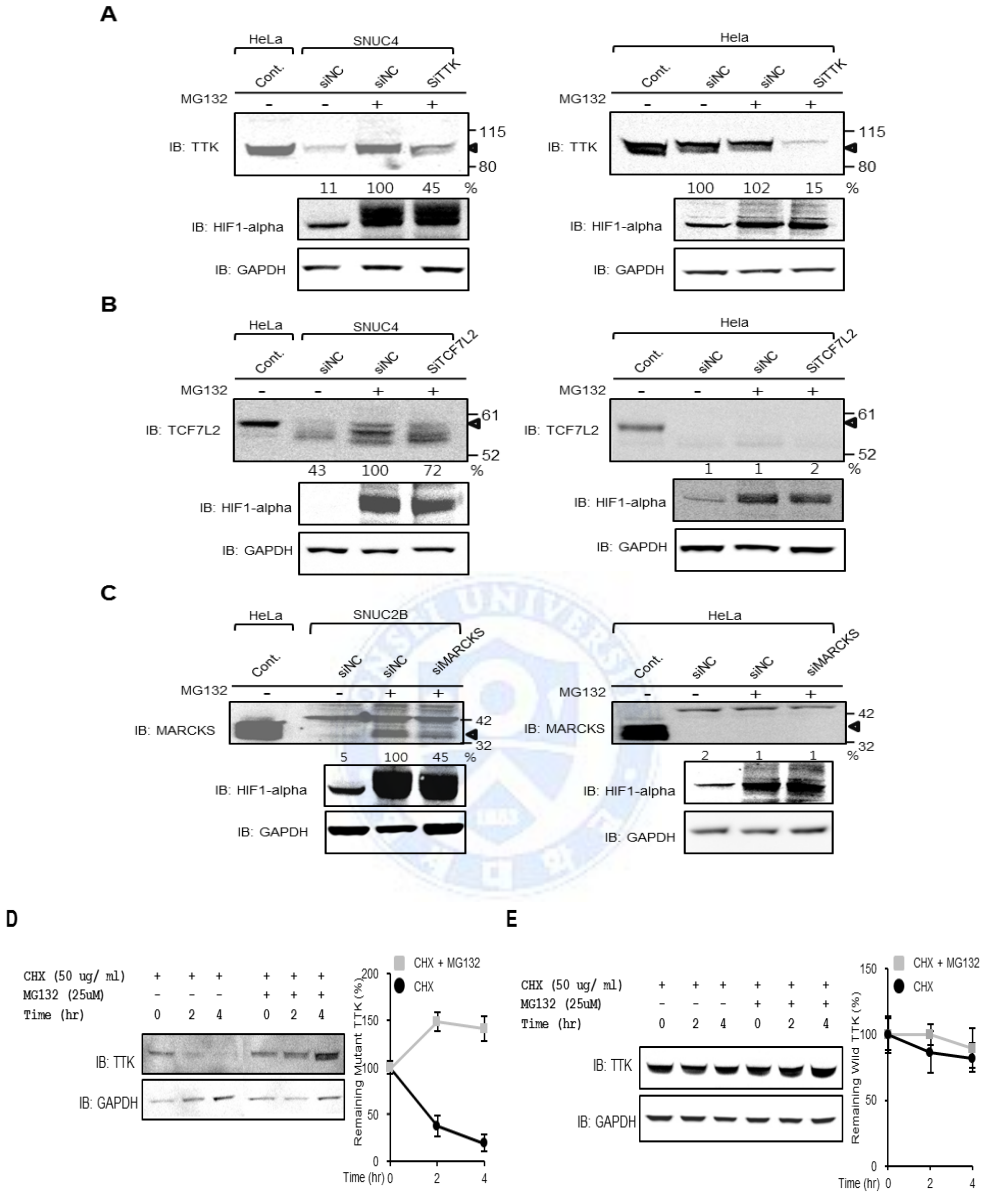


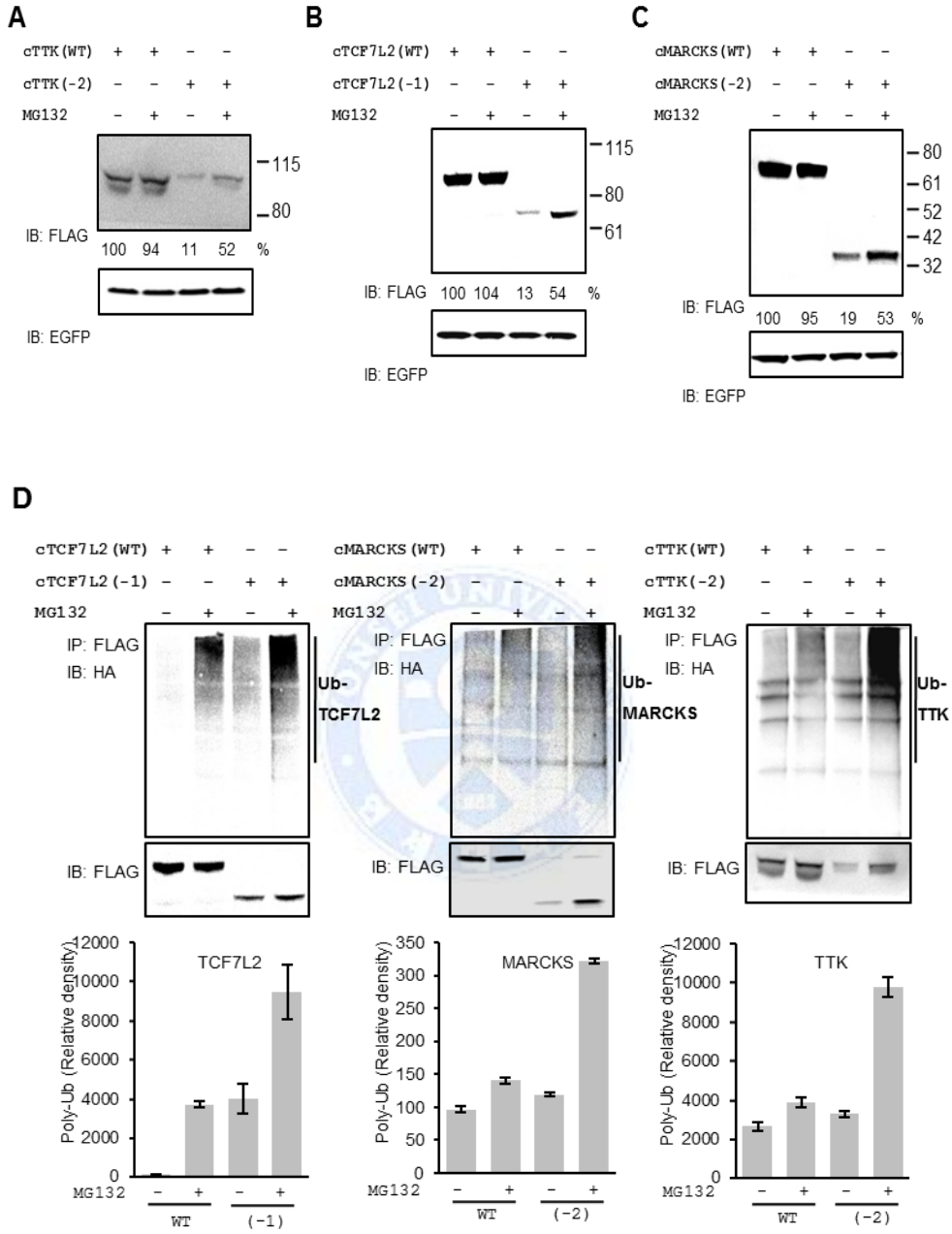

**Figure 7. Endogenous mutant TTK, TCF7L2, and MARCKS are detected after proteasome inhibition. (**A) Mutant TTK expression was evaluated in SNUC4 cells with homozygous *TTK* mutations. The increase in mutant TTK expression induced by MG132 treatment was decreased almost by 50% upon the transfection of siRNA against *TTK*. The same experiment was conducted in HeLa cells expressing wild-type TTK, and no significant

change in TTK expression was detected after MG132 treatment. (B) Mutant TCF7L2 expression was demonstrated in SNUC4 cells with homozygous mutations in *TCF7L2*. After MG132 treatment, a distinct band of similar size to the band detected in the control was detected, and this band disappeared after transfection of siRNA against *TCF7L2*. The same experiment was conducted in HeLa cells, but no mutant was detected after MG132 treatment. (C) Mutant MARCKS expression was evaluated in SNUC2B cells with heterozygous mutations in *MARCKS*. After MG132 treatment, a dramatic increase in the intensity of a band of similar size to the band corresponding to mutant MARCKS in the control was observed, and the intensity of this band was decreased by approximately 50% upon the transfection of siRNA against *MARCKS*. The same experiment was performed using HeLa cells, but no mutant protein was detected irrespective of MG132 or siRNA treatment. The controls used in these experiments were prepared from HeLa cells transfected with the cTTK(−2), cTCF7L2(−1), and cMARCKS(−2) vector constructs. Hif1α was used to confirm the consistent efficacy of MG132. To detect mutant MARCKS, an antibody specifically recognizing mutant MARCKS was used. (D) To measure the stability of mutant TTK, SNUC4 cells were incubated with cycloheximide (CHX) in the absence or presence of MG132. The expression of the TTK mutant was dramatically decreased by approximately 4-fold in the absence of MG132, but its expression was stable in the presence of MG132. (E) To compare the stability between mutant and wild-type TTK, the same experiment was carried out in HeLa cells expressing wild-type TTK. Wild-type TTK did not display any changes in protein expression upon cycloheximide treatment irrespective of the presence of MG132.

### 7. Overexpression of mutant MARCKS, TCF7L2, and TTK leads to heavy ubiquitination and localization around centrosomes

In addition to the rescue assays for endogenous mutant proteins, I further showed that mutant TTK, TCF7L2, and MARCKS are more unstable than their wild-type counterparts and rapidly degraded by the proteasome system when wild type and mutant TTK, TCF7L2, and MARCKS constructs were overexpressed in vitro. As expected, the mutant TTK, TCF7L2, and MARCKS protein expression levels were

approximately 11%, 13%, and 19%, respectively, of their wild-type levels. Inhibition of proteasome degradation by MG132 increased the expression of the mutants by approximately 3-fold (Fig. 8A–8C). To confirm the involvement of the proteasome system, I conducted an ubiquitination assay using the mutant and wild-type cDNA constructs of TTK, TCF7L2, and MARCKS. Consequently, the 3 mutant proteins were more heavily ubiquitinated (more than 2-fold) than the wild-type proteins in the presence of MG132 (Fig. 8D). On the basis of the ubiquitination assay, I further examined the localization of mutant proteins using immunofluorescence microscopy. According to previous studies, I hypothesized that if mutant proteins are specifically localized around centrosomes, which recruit the proteasomal machinery, then this further confirms that mutant proteins are actively degraded.[35,36] I conducted immunofluorescence staining using antibodies against FLAG and γ-tubulin, the latter of which was used as a centrosome marker. I chose mutant MARCKS constructs for the experiment. Under confocal microscopic examination, proteasome inhibition resulted in centrosomal expansion in cells transfected with both wild-type and mutant MARCKS constructs. However, colocalization with γ-tubulin was detected only in cells transfected with the mutant construct, indicating that mutant proteins are actively recruited to centrosomes after proteasome inhibition (Fig. 8E). Taken together, my results suggest that both endogenous and ectopically expressed mutant proteins are more rapidly degraded than wild-type proteins, and this process is mediated via the ubiquitin–proteasome pathway.


A
cTTK(WT) + + − −
cTTK(-2) − − + +
MG132 − + − +
115
80
IB: FLAG
100 94 11 52 %
IB: EGFP
B
cTCF7L2(WT) + + − −
cTCF7L2(-1) − − + +
MG132 − + − +
115
80
61
IB: FLAG 100 104 13 54 %
IB: EGFP
C
cMARCKS(WT) + + − −
cMARCKS(-2) − − + +
MG132 − + − +
80
61
52
42
32
IB: FLAG
100 95 19 53 %
IB: EGFP
D
cTCF7L2(WT) + + − −
cTCF7L2(-1) − − + +
MG132 − + − +
IP: FLAG
IB: HA
Ub-TCF7L2
IB: FLAG
cMARCKS(WT) + + − −
cMARCKS(-2) − − + +
MG132 − + − +
IP: FLAG
IB: HA
Ub-MARCKS
IB: FLAG
cTTK(WT) + + − −
cTTK(-2) − − + +
MG132 − + − +
IP: FLAG
IB: HA
Ub-TTK
IB: FLAG
TCF7L2
Poly-Ub (Relative density)
0 2000 4000 6000 8000 10000 12000
MG132 − + − +
WT (-1)
MARCKS
Poly-Ub (Relative density)
0 50 100 150 200 250 300 350
MG132 − + − +
WT (-2)
TTK
Poly-Ub (Relative density)
0 2000 4000 6000 8000 10000 12000
MG132 − + − +
WT (-2)

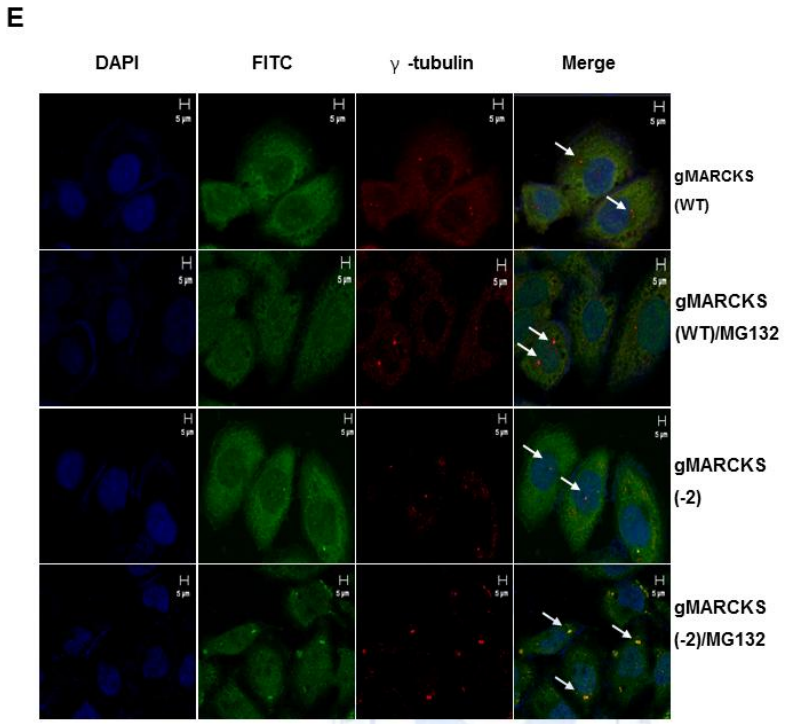


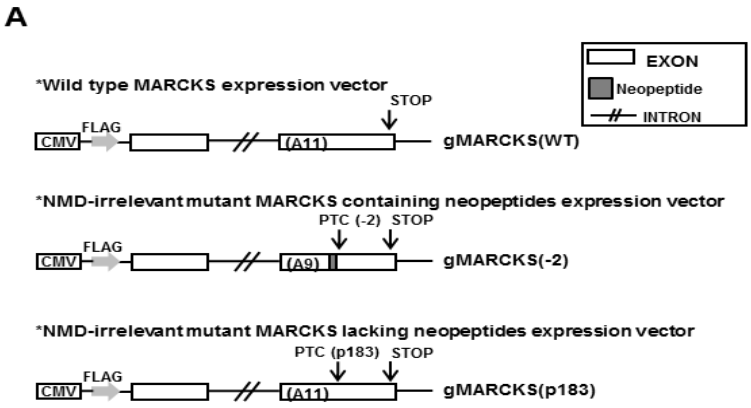

**Figure 8. Heavy ubiquitination of ectopically expressed mutant TTK, TCF7L2, and MARCKS and their colocalization with centrosomes after proteasome inhibition.** (A-C) Expression of wild-type and mutant TTK, TCF7L2, and MARCKS generated from the respective cDNA expression vectors before and after proteasome inhibition. Wild-type TTK, TCF7L2, and MARCKS expressions were much greater than those of their mutants, whereas increased protein expressions after proteasome inhibition were only detected for mutant proteins. (D) Ubiquitination assays of ectopically expressed wild-type and mutant TTK, TCF7L2, and MARCKS generated from cDNA expression vectors in the presence or absence of MG132. Mutant TTK, TCF7L2, and MARCKS were more heavily ubiquitinated than their wild-type counterparts (upper panel). The relative intensities of smear bands were quantified and presented as bar graphs (lower panel). (E) Subcellular localization of mutant MARCKS proteins generated from the cDNA expression vectors. Under confocal microscopic examination, proteasome inhibition resulted in centrosomal expansion in cells transfected with both wild-type and mutant MARCKS constructs. Colocalization with γ-tubulin was detected only in cells transfected with the mutant constructs, indicating that

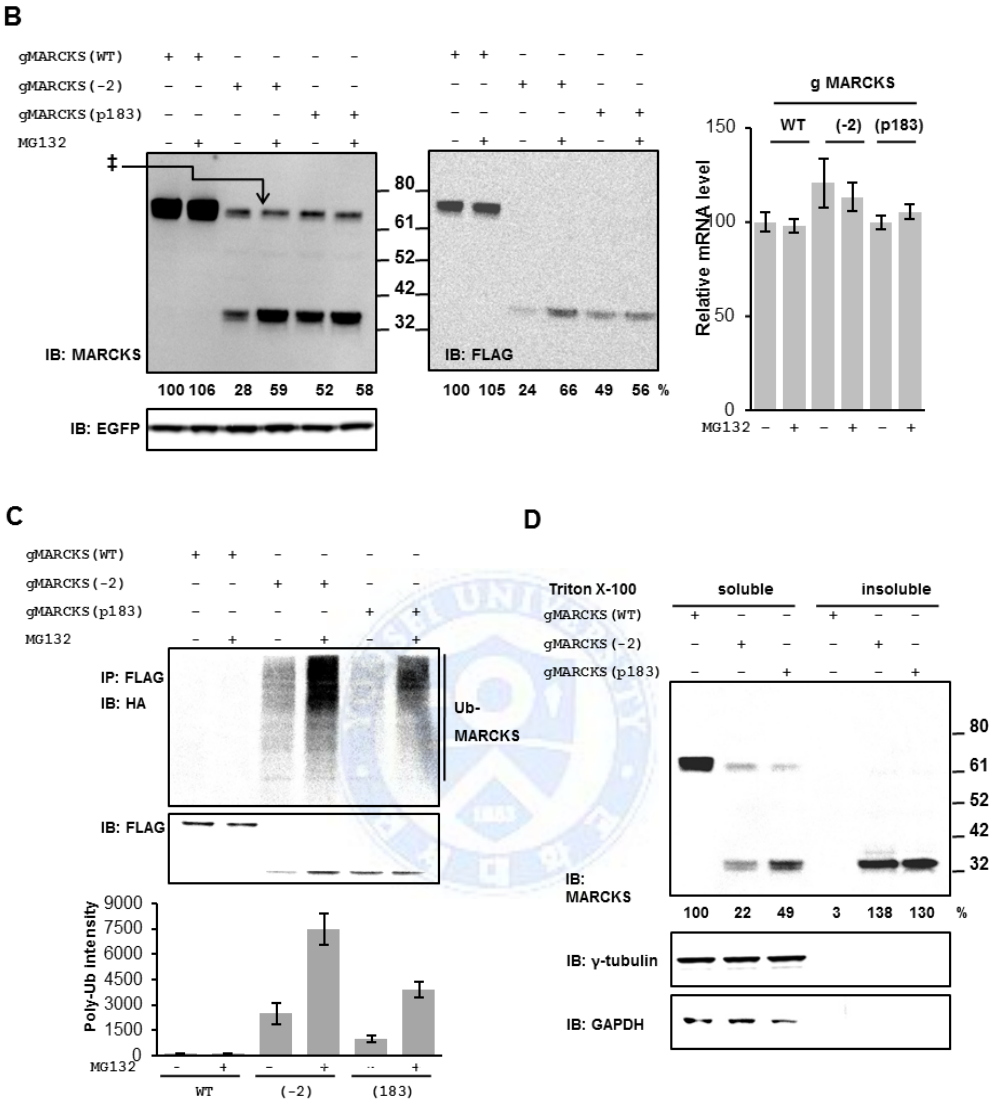
mutant proteins are actively recruited to the centrosomal region after proteasome inhibition. White arrows denote centrosomes.

## 8. Enhanced degradation of the truncated mutant MARCKS proteins containing neopeptides

All genes containing frameshift mutations in cMNRs in the last exon are expected to generate truncated proteins with variable lengths of neopeptides. After I showed that the endogenous neopeptide-containing truncated mutant proteins were mostly degraded in the proteasome, I hypothesized that neopeptides and/or protein truncation might enhance mutant protein degradation. To validate my hypothesis, I constructed another protein expression vector by using the MARCKS genomic DNA construct. The mutant protein produced from the gMARCKS(-2) vector was composed of 183 amino acids, 27 of which comprised neopeptides in the C-terminal region. I generated an additional vector construct by introducing an abnormal stop codon at amino acid position 183 [gMARCKS(p183)], which leads to generation of truncated mutant proteins lacking neopeptides (Fig. 9A). I evaluated MARCKS mRNA and protein expression levels before and after MG132 treatment. The expression of mRNAs from the cells transfected with the gMARCKS(-2) and gMARCKS(p183) vectors was similar to that in the cells transfected with the wild-type MARCKS vector [gMARCKS(WT); Fig. 9B, right]. Comparing the protein expression level, I found that the expression of ectopically expressed wild-type MARCKS protein was higher than that of ectopically expressed mutant MARCKS proteins irrespective of the presence of neopeptides. Interestingly, cells transfected with gMARCKS(p183) exhibited nearly 2-fold higher mutant protein levels than cells transfected with gMARCKS(-2) in the absence of MG132. After proteasome inhibition, the levels of neopeptide-containing mutant MARCKS were almost doubled, whereas those of the mutant lacking neopeptides were only increased slightly (Fig. 9B, left). The ubiquitination assay indicated that the mutant MARCKS

containing neopeptides is more heavily ubiquitinated than the mutant MARCKS lacking neopeptides, indicating that neopeptides are primarily responsible for the degradation of mutant MARCKS (Fig. 9C). As the expression of mutant MARCKS remained lower than that of wild-type MARCKS even after proteasome inhibition, I suspected that other factors might be involved in the low expression of mutant MARCKS. Thus, I hypothesized that the changes in mutant protein levels could be related to increases in insolubility, which contributes to the formation of insoluble bodies in cells. To show whether the formation of insoluble bodies reduces mutant protein expression, I fractionated the cell lysates into Triton X-100–soluble and Triton X-100–insoluble fractions. GAPDH and γ-tubulin were used as Triton X-100–soluble markers. Surprisingly, significant levels of both mutant MARCKS proteins were detected in the Triton X-100–insoluble fraction, indicating that the low expression of mutant MARCKS after proteasome inhibition was due to the increased insolubility of the mutant proteins and subsequent formation of insoluble bodies (Fig. 9D). These findings indicate that truncated mutant proteins containing neopeptides are rarely detected because of extensive degradation and increased insolubility (Fig. 9E).




B
gMARCKS(WT)
gMARCKS(-2)
gMARCKS(p183)
MG132
IB: MARCKS
IB: EGFP
IB: FLAG
100 106 28 59 52 58
100 105 24 66 49 56 %
g MARCKS
WT
(-2)
(p183)
Relative mRNA level
MG132
C
IP: FLAG
IB: HA
Ub-MARCKS
IB: FLAG
Poly-Ub intensity
WT
(-2)
(183)
D
Triton X-100
soluble
insoluble
IB: MARCKS
IB: γ-tubulin
IB: GAPDH
100 22 49 3 138 130 %

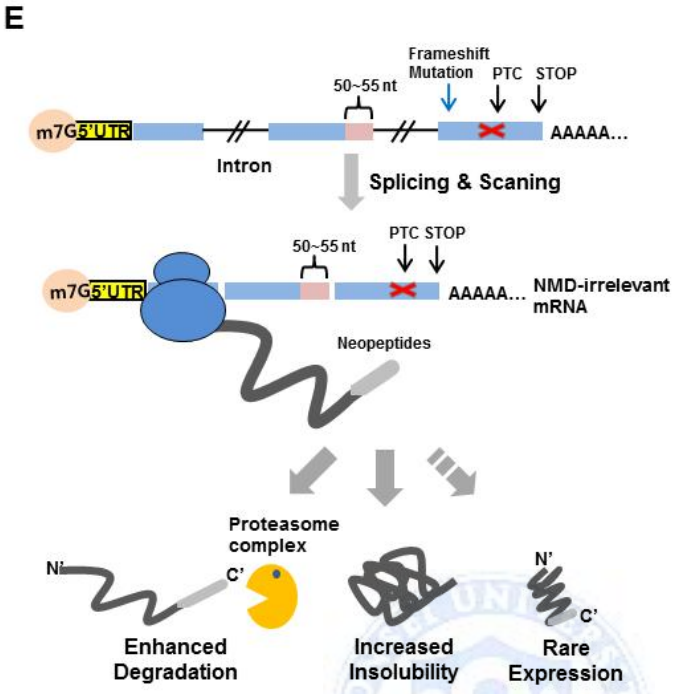



**Figure 9. Expression, degradation, and insolubility of wild-type and truncated mutant MARCKS containing or lacking neopeptides.** (A) Schematic diagram of the genomic DNA vector constructs of wild-type and mutant *MARCKS* constructs. (B) Protein expression from each vector construct was analyzed by Western blotting. The expression level of wild-type MARCKS was higher than that of both mutant MARCKS proteins, and the protein expression level of the mutant MARCKS lacking neopeptides was approximately 2-fold higher than that of the mutant MARCKS containing neopeptides (left panel). The mRNA expression level was also measured, and all of the constructs were expressed at similar levels (right panel). (C) An ubiquitination assay was conducted. Mutant MARCKS containing neopeptides was more heavily ubiquitinated, in contrast to the slight levels of ubiquitination of mutant MARCKS lacking neopeptides (upper panel). Relative intensities were measured and presented as a bar graph (lower panel). (D) Lysates from cells transfected with MARCKS(WT), MARCKS(−2), and MARCKS(183) were prepared from Triton-soluble supernatant and Triton-insoluble pellet fractions and analyzed by Western blotting. Wild-type MARCKS was mostly present in the soluble faction. Conversely, both mutant MARCKS proteins were mostly present in the insoluble faction. (E) Schematic

model for the fate of mutant proteins derived from PTC-containing mRNAs. ‡ denotes endogenous wild-type MARCKS.

### 9. Confirmation of the existence of NMD-resistant PTC-containing mutant *β-globin* mRNAs from gBglo-P39 and gBglo-P66 constructs in the steady-state

After I found NMD-irrelevant PTC-containing mRNAs are translated as efficiently as wild type mRNAs, I tried to study the mutant protein expressions from NMD-resistant and rescued PTC-containing mRNAs. Genomic human β-globin expression vectors with nonsense mutations were constructed in order to avoid enhanced degradation of neopeptide-containing mutant proteins. Using genomic DNAs from HeLa cells, the *β-globin* gene including exons and introns between the exons was cloned into a PCMV-3X-FLAG vector (gBglo-WT). Using the gBglo-WT construct, gBglo-P39 and gBglo-P66 constructs were generated by substitution mutagenesis generating PTCs at each 39$^{th}$ and 66$^{th}$ amino acid of a β-globin protein and these PTCs were expected to be recognized by NMD system (NMD-competent). In addition, gBglo-P101 and gBglo-P127 vectors were constructed by substitution mutagenesis generating PTCs at each 101$^{st}$ and 127$^{th}$ amino acid of a β-globin protein and these PTC locations were expected not to be recognized by NMD (NMD-irrelevant) (Fig. 10A). Two days after transfecting the β-globin expression constructs into HeLa cells, I harvested the cells and conducted RT-PCR to compare the expression level of precursor *β-globin* RNAs from each construct and qRT-PCR to compare the expression level of mature *β-globin* RNAs from each construct. Irrespective of the mutation status, the expression level of precursor *β-globin* RNAs from all constructs was almost similar and this result shows that the expressional variation among the five β-globin constructs are minimal. Regarding the expression level of mature *β-globin* RNAs, no significant difference was observed among the mature *β-globin* RNAs expression levels from gBglo-WT, gBglo-P101 and gBglo-P127 constructs, which shows NMD could not recognize the mRNAs from gBglo-

P101 and gBglo-P127 as substrates. On the other hand, the mRNA expression levels from gBglo-P39 and gBglo-P66 were about 30% of gBglo-WT mRNA expression level, which indicates NMD system is capable of removing only 70% of PTC-containing mRNAs from gBglo-P39 and gBGlo-P66 after splicing (Fig. 10B). These results suggest that constantly detected 30% of PTC-containing mRNAs are resistant to NMD and might be stable enough to generate mutant proteins in cells.

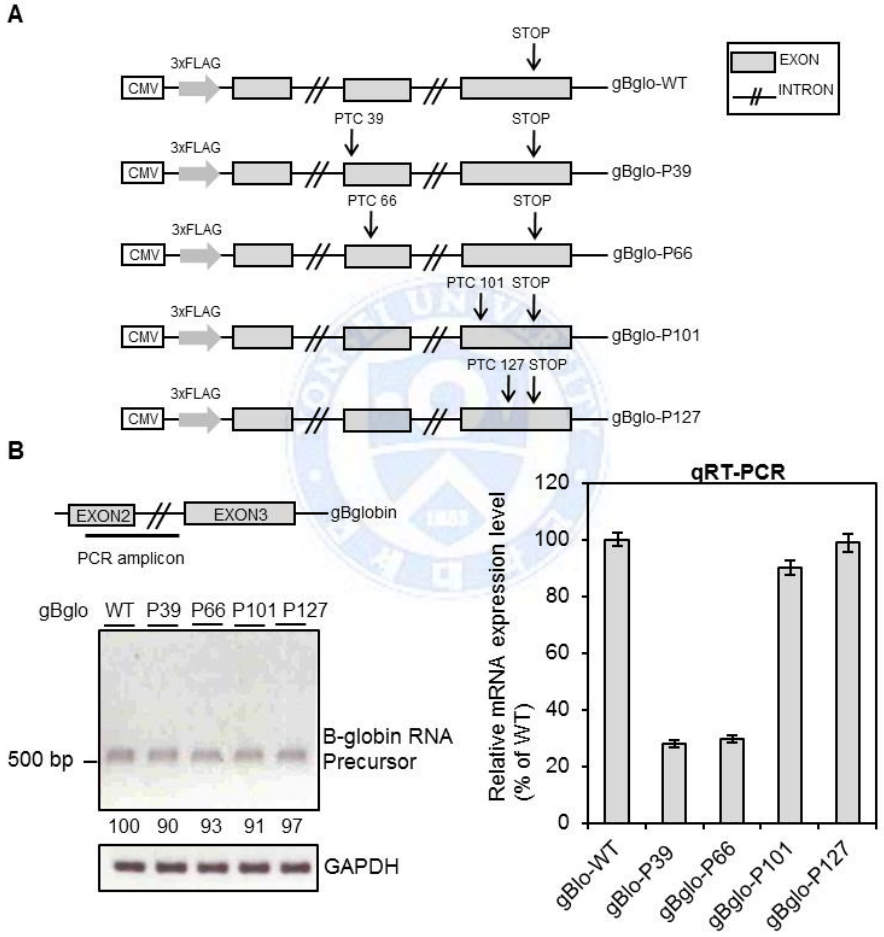



**Figure 10. Precursor and mature RNA expressions from constructed human β-globin vectors.** (A) The human genomic *β-globin* gene was cloned into a 3xFLAG expression vector (gBglo-WT) and subsequent mutagenesis was performed to generate mutant constructs (gBglo-P39, gBglo-P66, gBglo-P101 and gBglo-P127). (B) Precursor RNA

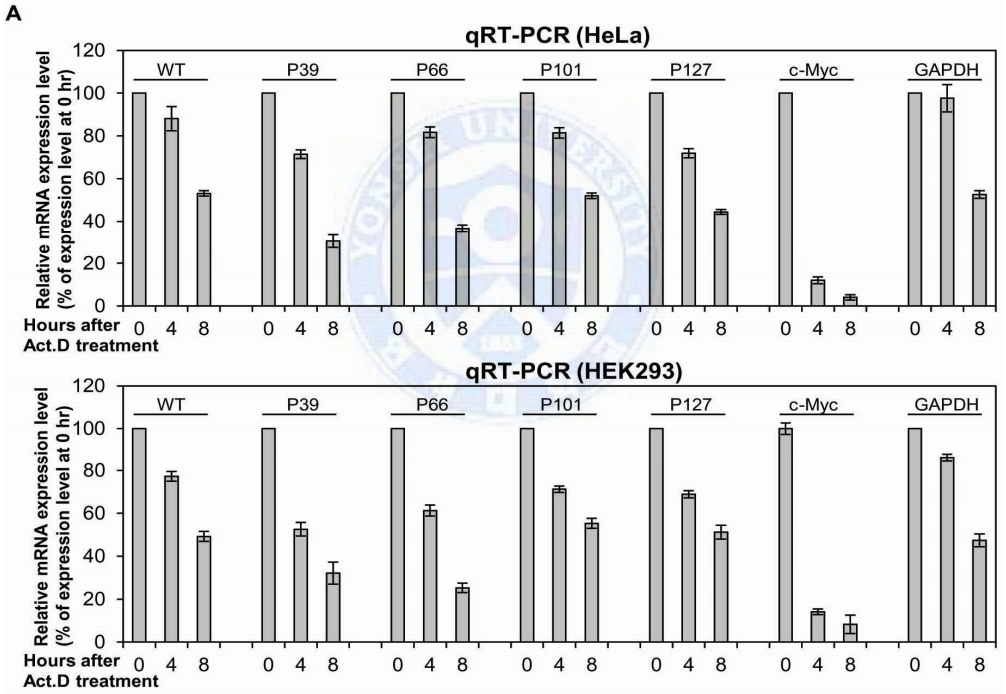
expression level was measured by RT-PCR and mature RNA expression level was measured by qRT-PCR. Precursor RNA expression levels from each gBglo construct were similar. Mature RNA expression levels from gBglo-WT, gBglo-P101 and gBglo-127 were similar and mRNA expression levels from gBglo-P39 and gBglo-P66 were about 30% of the expression level of gBglo-WT

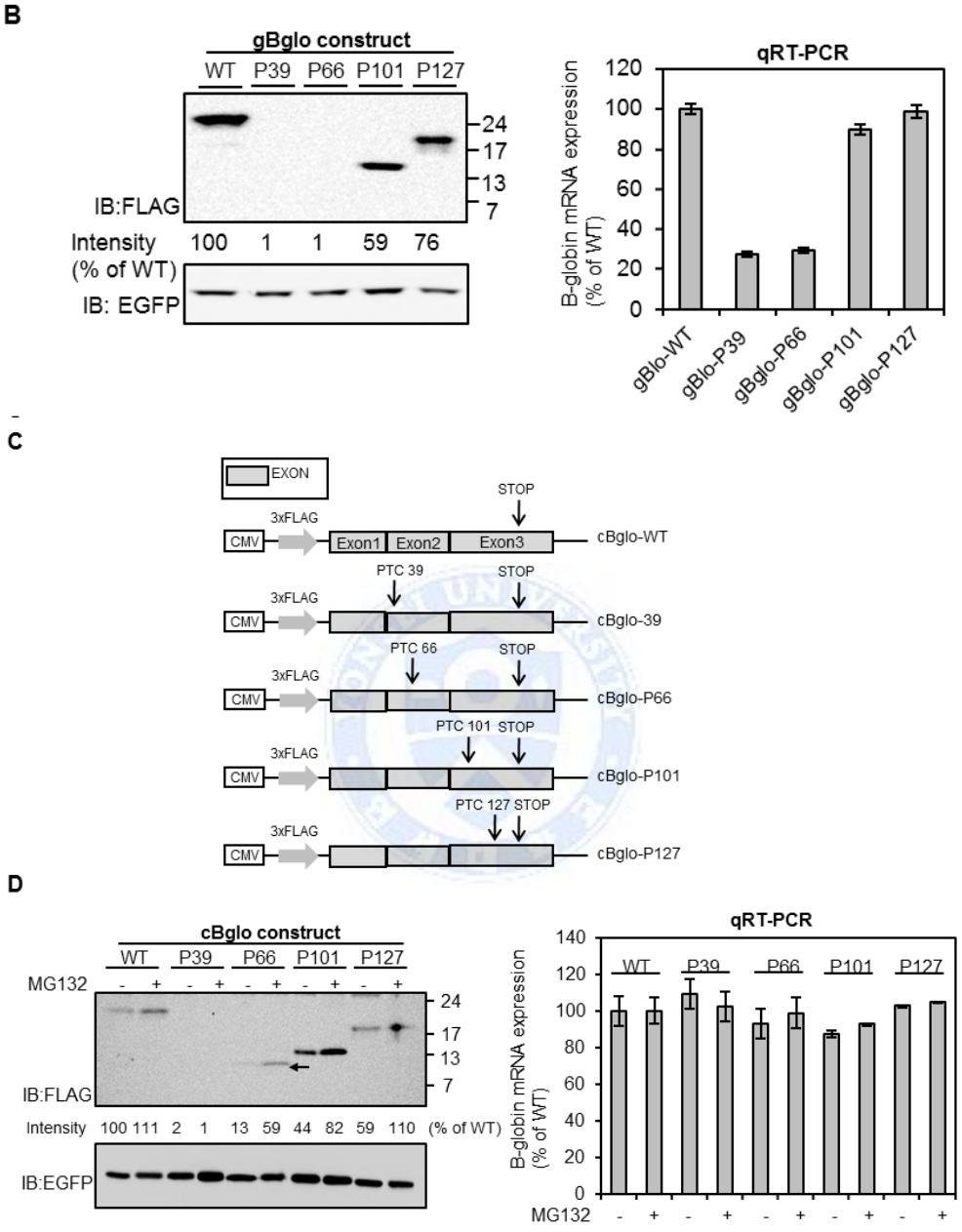
## 10. NMD-resistant PTC-containing mRNAs from gBglo-P39 and gBglo-P66 constructs are as stable as their PTC free counterparts but, mutant proteins are barely detected

After I have confirmed that the NMD-resistant mRNAs from gBglo-P39 and gBglo-P66 were constantly detected in cells, I tried to demonstrate whether mutant proteins are generated from these mRNAs. To be the proper sources for the generation of mutant proteins, mRNAs should be relatively stable in cells. Therefore, I analyzed the stability of mRNAs from each gBglo construct in two different cell lines. The gBglo constructs were transfected into both HeLa cells and HEK293 cells. And after 2 days of transfection, cells were treated with Actinomycin D (Act.D) for indicated time points before harvest. By performing qRT-PCR analysis, I found that the decreasing patterns of mRNA expression from the gBglo constructs at each 4 and 8 hours after Act.D treatment were quite similar, irrespective of mutation status.[21] Importantly, the NMD-resistant mRNAs from gBglo-P39 and gBglo-P66 were as stable as the mRNAs from both gBglo-WT and NMD-irrelevant gBglo-P101 and gBglo-P127 constructs. *GAPDH* mRNAs were used as a stable mRNA control and *c-Myc* mRNAs were used as a highly unstable mRNA control.[37] The mRNA expression levels from gBglo constructs were shown as relative values to each mRNA expression level before Act.D treatment (Fig. 11A). As the NMD-resistant mRNAs were relatively stable even after transcriptional inhibition, I expected that mutant proteins might be generated from these mRNAs. To confirm my hypothesis, the five gBglo constructs were transfected into HeLa cells and transfected cells were harvested after 2 days. Then

both western blotting and qRT-PCR were performed. The results showed that the expression levels of mutant proteins from gBglo-P101 and gBglo-P127 (NMD-irrelevant constructs) were 62 % and 73 % of wild type β-globin expression level from gBglo-WT. However, no mutant proteins were detected in cells transfected with gBglo-P39 and gBglo-P66 constructs (Fig. 11B). After I found lack of mutant protein expressions from NMD-competent constructs (gBglo-P39 and gBglo-P66), I generated completely NMD-irrelevant cDNA β-globin expression constructs (cBglo-WT, cBglo-P39, cBglo-P66, cBglo-P101 and cBglo-P127) to determine the objective stabilities of mutant β-globin proteins first (Fig. 11C). I transfected the five cDNA β-globin constructs into HeLa cells and this time, cells were treated with MG132 to prevent the potential degradation of mutant proteins. Expressions of proteins and mRNAs were analyzed by western blotting and qRT-PCR. The mutant β-globin proteins from cBglo-P101 and cBglo-P127 were stably detected and significant increases of these mutant proteins were observed when MG132 was treated. On the other hand, mutant β-globin proteins from cBglo-P39 were not detected even after MG132 treatment, which shows that the mutant proteins from cBglo-P39 are extremely unstable and the gBglo-P39 construct is not a proper model to study the generation of mutant proteins from PTC-containing mRNAs. Importantly, small amount of mutant β-globin proteins from cBglo-P66 was detected and the mutant protein expression was significantly increased up to 60 % of wild type β-globin protein expression after MG132 treatment. This result indicates that the mutant proteins from the cBglo-P66 construct are stable enough to be detected and using the gBglo-P66 construct, generation of mutant proteins from NMD-resistant mRNAs can be evaluated when MG132 is treated. Taken together, I demonstrated that the gBglo-P66 construct was a proper model to study the generation of mutant proteins from NMD-resistant mRNAs (Fig. 11D). Then, using genomic β-globin expression vectors, I conducted the same experiment that had been performed with cDNA β-globin expression vectors. The results showed that β-globin mutant proteins from NMD-irrelevant constructs gBglo-P101 and gBglo-

P127 were relatively stable and increased after MG132 treatment. On the other hand, mutant proteins from gBglo-P39 and gBglo-P66 were barely detected and only after MG132 treatment, the expression levels of β-globin mutant proteins were slightly increased up to 12% and 17% of the expression level of β-globin proteins from gBglo-WT. Based on these results, I calculated the efficiency of protein generation by dividing the protein expression level with mRNA expression level when MG132 is treated. As a result, I found that mutant proteins from gBglo-P39 and gBglo-P66 were very inefficiently generated, compared to gBglo-WT, gBglo-P101 and gBglo-P127 (Fig. 11E and 11F).




B
gBglo construct
WT P39 P66 P101 P127
24
17
13
7
IB:FLAG
Intensity 100 1 1 59 76
(% of WT)
IB: EGFP
qRT-PCR
B-globin mRNA expression (% of WT)
120
100
80
60
40
20
0
gBlo-WT
gBlo-P39
gBglo-P66
gBglo-P101
gBglo-P127
C
EXON
STOP
3xFLAG
CMV
Exon1 Exon2 Exon3
cBglo-WT
PTC 39
cBglo-39
PTC 66
cBglo-P66
PTC 101
cBglo-P101
PTC 127
cBglo-P127
D
cBglo construct
WT P39 P66 P101 P127
MG132 - + - + - + - + - +
24
17
13
7
IB:FLAG
Intensity 100 111 2 1 13 59 44 82 59 110 (% of WT)
IB:EGFP
qRT-PCR
B-globin mRNA expression (% of WT)
140
120
100
80
60
40
20
0
WT P39 P66 P101 P127
MG132 - + - + - + - + - +

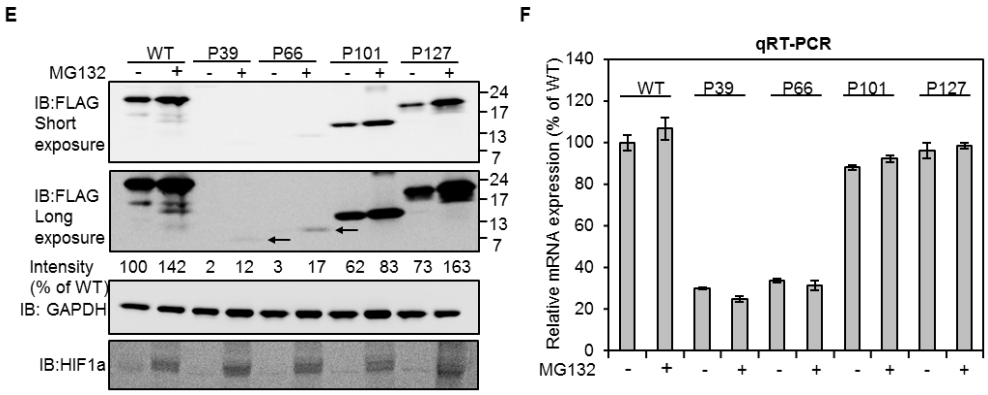



**Figure 11. PTC-containing mRNAs from gBglo-P39 and gBglo-P66 exist stably, but barely generate mutant proteins.** (A) HeLa cells were transfected with gBglo-WT, gBglo-P39, gBglo-P66, gBglo-P101 and gBglo-P127 and treated with actinomycin D before harvest. Cells were harvested at 0 hr, 4 hrours and 8 hrours after actinomycin D treatment and qRT-PCR analysis was performed to measure the mRNA expression levels at each time points. The mRNA expression level is shown as relative value to the level at 0 hr. The decreasing patterns of mRNA expression levels from gBglo constructs were similar. (B) Genomic human β-globin expression vectors were transfected into HeLa cells. Western blotting and qRT-PCR were performed to evaluate the expressions of proteins and mRNAs. Stable mRNA and proteins expressions from gBglo-WT, gBglo-P101 and gBglo-P127 were observed, but no protein expressions were observed from gBglo-P39 and gBglo-P66. Low expression level of PTC-containing mRNAs from gBglo-P39 and gBglo-P66 was detected. (C) To assess the objective stability of proteins from gBglo constructs, cDNA forms of human β-globin expression vectors were constructed. (D) After transfecting cBglo constructs into HeLa cells, protein and mRNAs expression levels were determined. The expression levels of mRNAs from cBglo constructs were almost similar and the proteins from cBglo-WT, cBglo-P101 and cBglo-P127 were stably expressed in cells. The protein expression level from cBglo-P39 was undetectable, but the proteins expression level from cBglo-P66 was significantly increased by proteasome inhibition (13 % to 60 % of wild type β-globin expression level). (E) Mutant protein expressions from gBglo constructs were determined by western blotting analysis. Only trace amount of mutant proteins was barely detected from gBglo-P66 after MG132 treament. (F) Using cells transfected with gBglo constructs, qRT-PCR was performed to measure relative mRNA expression levels from each

gBglo construct. MG132 treamtment showed no influence on the expression of mRNAs from gBglo constructs.

### 11. Overexpression of NMD-resistant PTC-containing mRNAs from the gBglo-P66 construct does not affect the generation of mutant proteins

After I found that mutant proteins were not efficiently generated from the NMD-resistant mutant mRNAs, I suspected that very low efficiency of the protein generation might be due to the low amount of the mutant mRNAs. Therefore, I gradually increased the plasmid amount of gBglo-P66 transfected into cells. Transfecting gBglo-WT and gBglo-P66 vectors into HeLa cells, I analyzed the mRNA expression levels as the amount of the transfected plasmid increased. The results showed that the ratio between the mRNA expression levels from gBglo-WT and gBglo-P66 was almost constant as 10:3 when the same amount of gBglo-WT and gBglo-P66 vectors were transfected. Interestingly, I found that NMD-resistant mRNAs from gBglo-P66 were accumulated up to 85 % of the mRNA expression level from gBglo-WT when three times more plasmids of gBglo-P66 than gBglo-WT were transfected. These results suggest that NMD have limited capacity to eliminate PTC-containing mRNAs in cells (Fig. 12A and 12B). Then, I tried to find whether the mutant protein expression was increased as the level of accumulated NMD-resistant mRNAs from gBglo-P66 was increased. HeLa cells were transfected with gBglo-WT and gradually increasing amount of gBglo-P66. After 2 days, cells were treated with MG132 and harvested to analyze both mRNA and protein expression levels. The results showed that the NMD-resistant mutant mRNA expression level was gradually increased according the amount of transfected gBglo-P66 plasmid was increased (Fig. 12C). However, I could not detect increase level of mutant proteins from accumulated mutant mRNAs from gBglo-P66 and the efficiency of mutant protein generation showed almost no change, which suggests that the NMD-resistant mutant mRNAs from gBglo-P66 are not translationally

active (Fig. 12C and 12D).

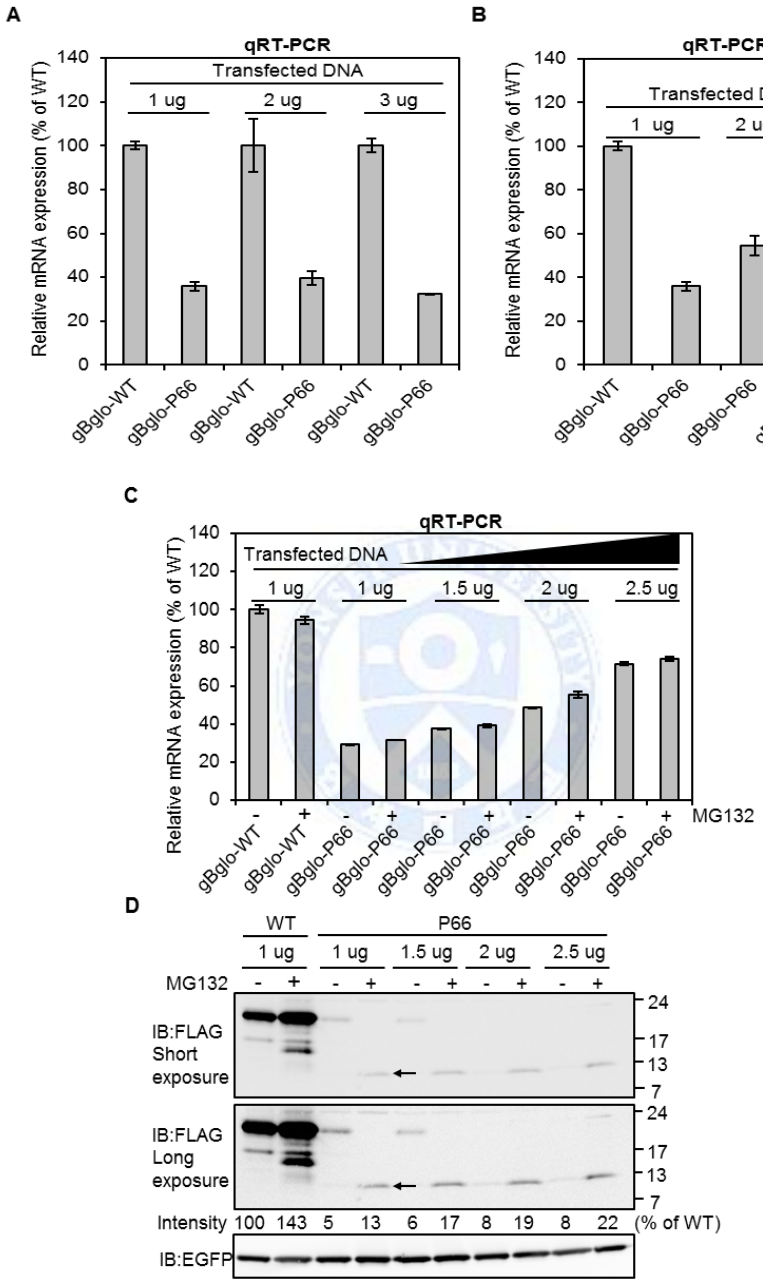



**Figure 12. Transfection of increasing gBglo-P66 plasmids leads to the accumulation of NMD-resistant mRNAs, but the generation of mutant proteins is not changed.** (A) Cells were transfected with gradually increasing amount of gBglo-P66 plasmids or constant

amount of gBglo-WT plasmids. Then, qRT-PCR was performed. The overexpression of gBglo-P66 caused accumulation of NMD-resistant mutant mRNAs from gBglo-P66. (B) Western blotting using the same samples used for qRT-PCR analysis was performed to evaluate the mutant protein expression levels from increased mutant mRNAs from gBglo-P66. The mutant protein expression level was barely increased, while the mutant mRNA expression level from gBglo-P66 was gradually increased.

## 12. Mutant proteins from NMD-resistant PTC-containing mRNAs from the gBglo-P66 construct were mainly generated in the pioneer round of translation

After splicing and export to the cytoplasm, mature RNAs usually subject to the pioneer round of translation and then, if the mRNAs have no defects, these mRNAs subject to the eIF4E dependent translation which is designated as bulky translation. It has been well known that most proteins in cells are generated by the bulky translation step, but the detection of small amount of proteins generated in the pioneer round of translation step had also been reported.[38,39] Based on these ideas, I tried to identify the sources of the trace amount of β-globin mutant proteins from the gBglo-P66 construct in cells. I suspected that the trace amount of mutant proteins from NMD-resistant mutant mRNAs might be generated from the pioneer round of translation. To determine the involvement of mutant proteins in the pioneer round of translation, I constructed a vector expressing 4E-BP1 which specifically inhibits eIF4E dependent translation.[40] Transfecting increasing amount of 4E-BP1 plasmids into HeLa cells with constant amount of EGFP control vector, I demonstrated that 4E-BP1 proteins effectively inhibited the bulky translation in cells (Fig. 13A). Then, I performed an experiment to determine whether the generation of mutant proteins from gBglo-P66 was affected by 4E-BP1 overexpression. HeLa cells were transfected with gBglo-WT or gBglo-P66 with or without 4E-BP1. The western blotting results showed that the amount of β-globin mutant proteins from gBglo-P66 was not reduced, while the amount of wild type β-globin proteins from gBglo-WT was significantly reduced as the amount of 4E-BP1

proteins increase, which shows that the β-globin mutant proteins from gBglo-P66 were rarely generated from bulky translation (Fig. 13B). I also confirmed this result by using another eIF4E specific translational inhibitor (4EGi). As a result, the β-globin mutant proteins from gBglo-P66 were also barely affected by 4EGi treatment while, the wild type β-globin proteins were significantly reduced after 4EGi treatment (Fig. 13C). Based on these results, I concluded that the mutant proteins from NMD-resistant mRNAs were mostly generated from the pioneer round of translation step.

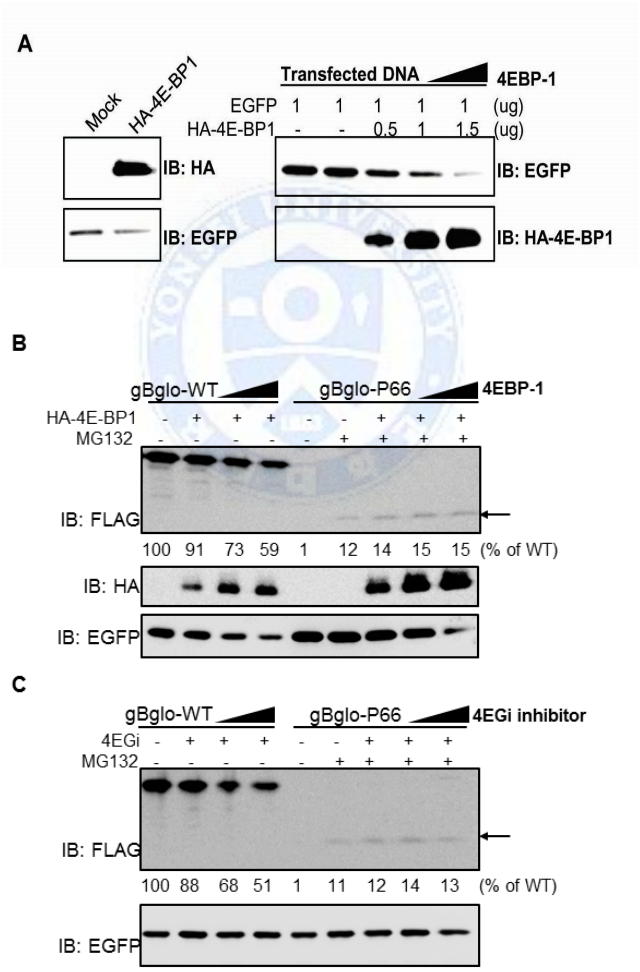



**Figure 13. Trace amount of mutant proteins from gBglo-P66 were mostly generated in the pioneer round of translation.** (A) A 4E-BP1 expression vector was constructed and

demonstrated to inhibit specifically eIF4E dependent translation (bulky translation). After HeLa cells were transfected with constant amount of an EGFP control plasmid and gradually increased amount of the 4E-BP1 expression plasmid, protein expression level of EGFP was analyzed by western blotting. EGFP protein expression was dramatically reduced as the 4E-BP1 expression level was increased. (B-C) To determine the involvement of the trace amount of mutant proteins from gBglo-P66 in bulky translation, gBglo-WT and gBglo-P66 were transfected into HeLa cells with increasing amount of 4E-BP1 plasmids. Cells were harvested and western blotting was performed. The expression level of mutant proteins from gBglo-P66 was not affected by 4E-BP1 or 4EGi, while the protein expression level was significantly reduced by bulky translational inhibition.

## 13. NMD-resistant PTC-containing mutant mRNAs from gBglo-P66 are not associated with polysomes

To assess the efficiency of translation of NMD-resistant mRNAs from gBglo-P66, I conducted polysome analysis using gBglo-P66, cBglo-P66, gBglo-WT and cBglo-WT constructs. HeLa cells were transfected with each construct with nontargeting siRNA (siNC) or siRNA targeting UPF1 (siUPF1). Three days after transfection, cells were harvested and cell lysates were fractionated according to the previously well-known polysome analysis methods. Twelve fractions were collected and subsequently, RNAs were extracted to perform RT-PCR. The results revealed that the *β-globin* mRNAs from gBglo-WT and cBglo-WT were mainly detected in polysome associated fractions (right shifted). Also *GAPDH* mRNAs were associated with polysome fractions (Fig. 14A and 14B). In case of gBglo-P66, the distribution pattern of *β-globin* mRNAs from gBglo-P66 was left shifted compared to *β-globin* mRNAs from cBglo-P66, which shows that NMD-resistant mutant mRNAs from gBglo-P66 are translationally repressed (Fig. 14C and 14D). Importantly, I found that the distribution of *β-globin* mRNAs from gBglo-P66 was right shifted when NMD was inhibited by down-regulating UPF1, a key NMD factor (Fig. 14E and 14F). Taken together, I conclude that 1) NMD-resistant mutant

mRNAs from gBglo-P66 are not actively translated in cells and 2) UPF1 plays important roles in the translational inhibition of the PTC-containing mRNAs.

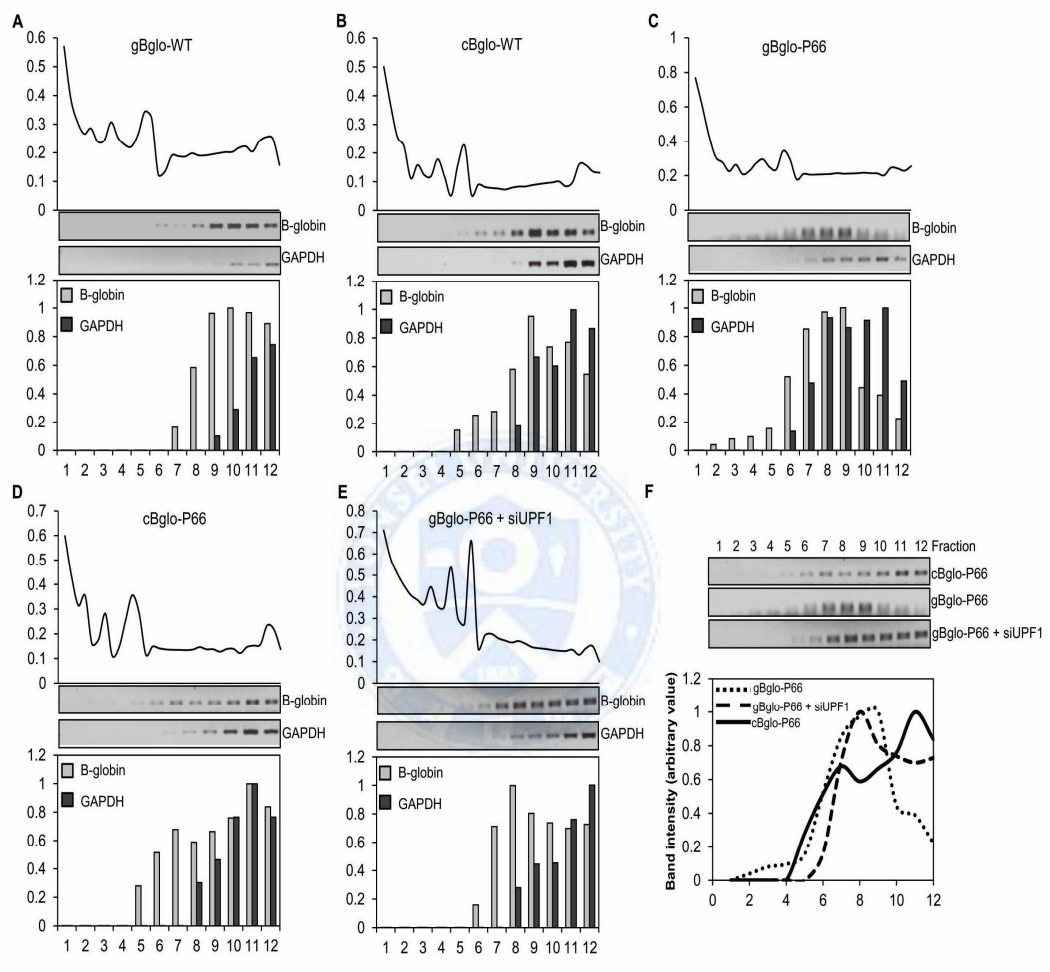



**Figure 14. NMD-resistant mutant mRNAs from gBglo-P66 are translationally repressed.** (A-B) After polysome fractionation, using RNAs extracted from collected fractions, RT-PCR was performed to evaluate the translational competence of mRNAs from gBglo-WT and cBglo-WT. *β-globin* mRNAs from both constructs were mostly detected in 8 to 12 fractions and *GAPDH* mRNAs were mainly detected in 9 to 12 fractions (right shifted). (C-D) After polysome fractionation, RT-PCR was also performed using RNAs from cells transfected with gBglo-P66, cBglo-P66. NMD-resistant mutant mRNAs from gBglo-P66

were mostly detected in 7 to 9 fractions, while mutant mRNAs from cBglo-P66 were mostly detected in 10 to 12 fractions. In both cells transfected with gBglo-P66 and cBglo-P66, *GAPDH* mRNAs were mostly found in 8 to 12 fractions. The distribution pattern of NMD-resistant mRNAs from gBglo-P66 was left shifted compared the pattern of mutant mRNAs from cBglo-P66. (E) Polysome associated mutant mRNAs from gBglo-P66 were detected when NMD was inhibited by UPF1 knock-down. (F) The comparison between the mutant mRNAs from gBglo-P66, cBglo-P66 and gBglo-P66 with UPF1 knock-down. NMD-resistant mRNAs from gBglo-P66 were translationally repressed, but the repression was relieved when NMD was inhibited by UPF1 knock-down**.**

## 14. Mutant proteins are actively translated from the rescued PTC-containing mRNAs of the gBglo-P66 construct when NMD is inhibited by UPF1 knock-down

Based on the protein expression analysis of NMD-resistant PTC-containing mRNAs, I wondered if PTC-containing mRNAs are rescued from NMD, the bulky translation of the rescued mutant mRNAs is also inhibited. To address this question, I first compared the expression level of mRNAs from cells transfected with gBglo-WT and from cells transfected with gBglo-P66 when NMD was inhibited or not. The qRT-PCR results indicated that the expression level of mutant *β-globin* mRNAs from gBglo-P66 was increased up to about 60 % of the expression level of mRNAs from gBglo-WT when NMD was inhibited by UPF1 knock-down (Fig. 15A). Then I analyzed the protein expression levels from gBglo-WT and gBglo-P66 when NMD was inhibited by UPF1 knock-down. The β-globin protein expression level from gBglo-WT was not affected by UPF1 down-regulation and the β-globin protein expression level was about 1.4 fold increased after MG132 treatment, which indicates that the protein expression level from gBglo-WT is barely affected by NMD inhibition. On the other hand, the expression level of mutant proteins from gBglo-P66 was increased more than three times when NMD was inhibited and MG132 was treated (Fig. 15B). Based on polysome analysis performed above, I

assumed that the increased level of mutant proteins might be generated from bulky translation. To confirm this idea, I measured the mutant protein expression levels from gBglo-P66 in several different conditions. Using 4E-BP1 and 4EGi inhibitor, bulky translation was blocked in conditions of NMD inhibition or NMD activation. The results showed that the expression level of β-globin mutant proteins from gBglo-P66 was significantly increased when NMD was inhibited by UPF1 knock-down, but the bulky translation inhibition by 4EGi or 4E-BP1 blocked the increase of the β-globin mutant protein expression from gBglo-P66, which indicates that the increased β-globin mutant proteins from gBglo-P66 are mainly generated by eIF4E dependent bulky translation (Fig. 15C).

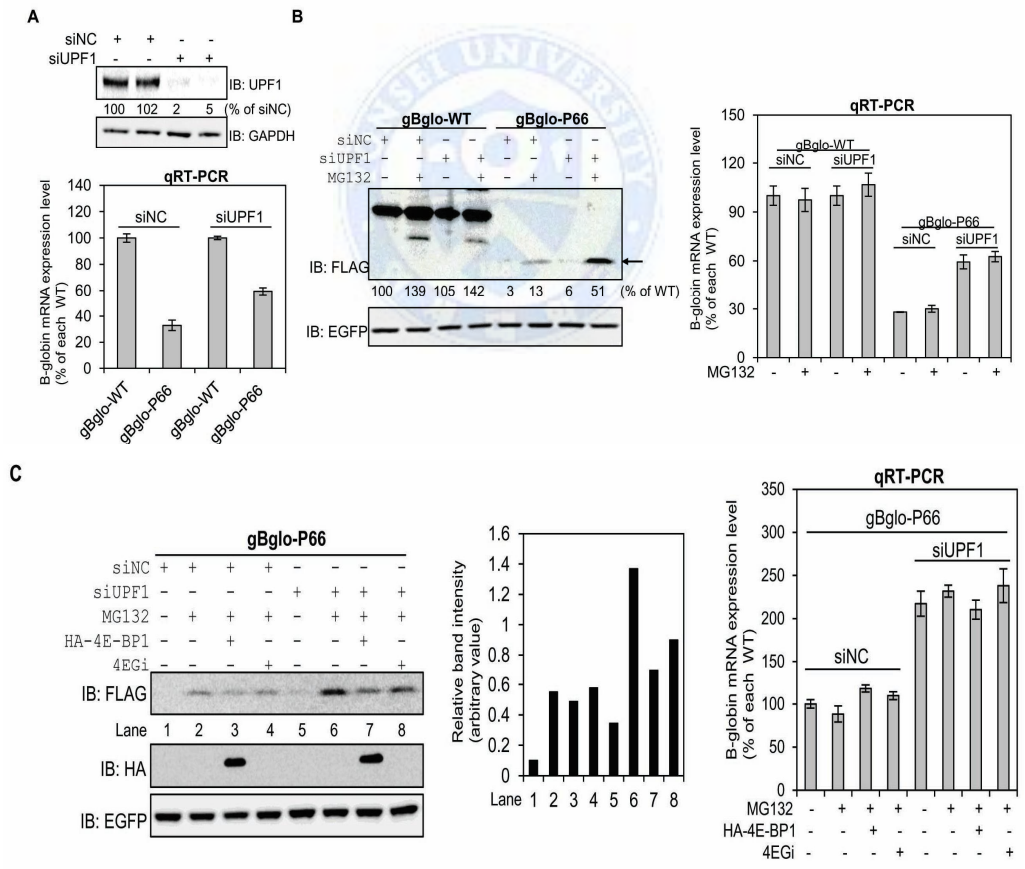



**Figure 15. Significantly increased amount of mutant proteins from gBglo-P66 was**

**detected when NMD was inhibited by UPF1 knock-down.** (A) The efficiency of RNAi against human *UPF1* was confirmed by transfecting siUPF1 into HeLa cells with gBglo-WT or gBglo-P66. After 3 days of transfection, western blotting and qRT-PCR were performed. UPF1 was significantly inhibited by siUPF1 transfection and the expression level of mutant mRNAs from gBglo-P66 almost doubled compared with the expression level of mutant mRNAs from gBglo-P66 with no NMD inhibition. (B) With or without UPF1 knock-down, the protein expression levels from gBglo-WT and gBglo-P66 were measured by western blotting. When UPF1 was down-regulated, slight increase of the protein expression level from gBglo-WT was observed, irrespective of MG132. By NMD inhibition, the protein expression level of gBglo-P66 was increased three times more than the mutant protein expression level from gBglo-P66 with no NMD inhibition. The mRNA expression levels were also confirmed by qRT-PCR. (C) To demonstrate the involvement of the large amount mutant protein generation from gBglo-P66 in bulky translation, 4E-BP1 and 4EGi were cotransfected or treated to inhibit bulky translation of mutant mRNAs from gBglo-P66. The increased expression level of mutant proteins from gBglo-P66, when NMD was inhibited, was decreased to almost the same protein expression level of gBglo-P66 with no NMD inhibition by bulky translational inhibition. By qRT-PCR analysis, the mRNA expression levels were measured and not affected by 4E-BP1 and 4EGi treatment.

## 15. UPF1 plays key roles in the translational regulation of PTC-containing mRNAs

After I have confirmed that a large amount of mutant proteins from the rescued PTC-containing mRNAs is generated from the bulky translation step, I further tried to demonstrate if NMD inhibition itself is enough for the generation of a large amount of mutant proteins or UPF1 plays another role besides NMD inhibition. Core NMD factors (UPF1 and SMG1) and core EJC factors (Y14 and EIF4A3) were selected to find out if the down-regulation of these four factors by RNAi would lead to the generation of a large amount of mutant proteins from the gBglo-P66 construct. After I confirmed that each NMD relevant factor was efficiently down-regulated by RNAi, I analyzed both the mRNA and protein expression levels

from gBglo-P66. The results showed that the increased mRNA expression levels from gBglo-P66 were almost similar (about 2-fold increases) when NMD was inhibited by each specific RNAi treatment, which indicates that NMD can be inhibited by down-regulating any of these four factors (Fig. 16A and 16B). However, when I measured the mutant protein expression levels from gBglo-P66, I observed that only down-regulations of UPF1 or SMG1 leaded to the significant increase of mutant protein expressions from gBglo-P66. No mutant protein increases were detected when NMD was inhibited by manipulating core EJC factors Y14 and EIF4A3 (Fig. 16C). Based on the fact that the known roles of SMG1 are to activate and regulate UPF1 activity by phosphorylation, I concluded that UPF1 plays key roles in the translational regulation of NMD-competent PTC-containing mRNAs.

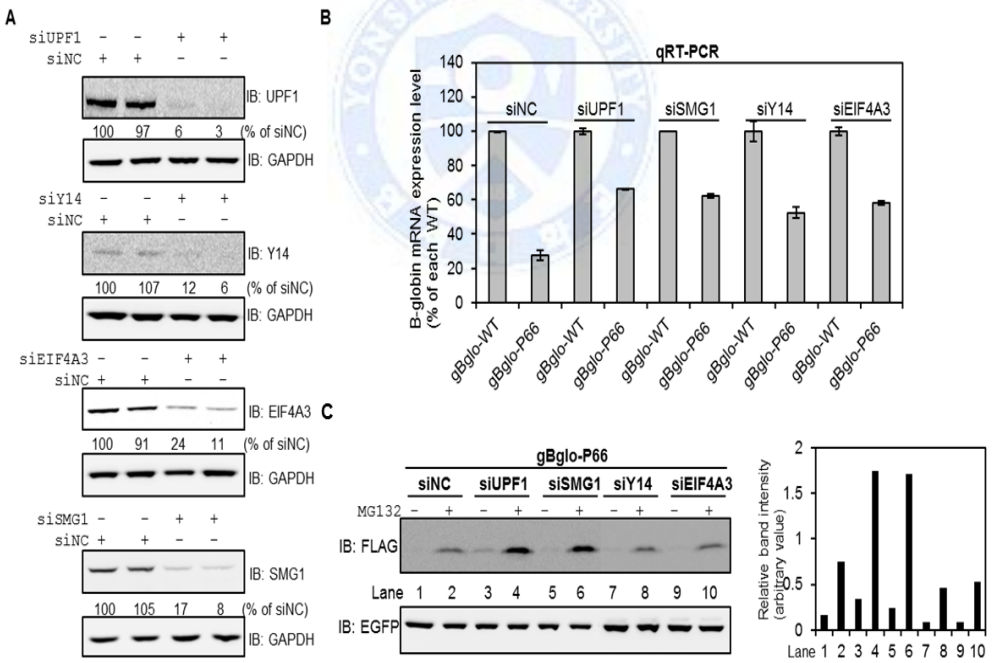



**Figure 16. UPF1 plays key roles for bulky translation of rescued PTC-containing mRNAs from the gBglo-P66 construct.** (A) To demonstrate the involvement of core NMD and EJC factors in the generation of mutant proteins, each specific RNAi was transfected

into HeLa cells and significant down-regulations of UPF1, SMG1, Y14 and EIF4A3 were confirmed. (B) The analysis of mRNA expression levels from the gBglo-P66 showed almost the same increases when NMD was inhibited by down-regulating each of 4 factors. (C) Significant increases (about 3-fold increases) of mutant protein expressions from the gBglo-P66 construct were only detected when UPF1 or SMG1 were down-regulated.

## 16. NMD bypassed PTC-containing mRNAs from the CrPV-gBglo-P66 construct are still translationally repressed by UPF1

After I found that UPF1 plays specific roles for the active generation of mutant proteins from rescued PTC-containing mRNAs, I further tried to demonstrate if UPF1 still affect the translational regulations of NMD bypassed PTC-containing mRNAs using CrPV-IRES constructs. It has been well known that NMD can be efficiently bypassed by CrPV-IRES elements which directly recruit ribosomes without any translational initiation complex proteins.[41] Therefore, I generated CrPV-gBglo-WT and CrPV-gBglo-P66 constructs by inserting CrPV-IRES sequences upstream of each genomic B-globin wild type and PTC66 constructs (Fig. 17A). Three days after transfecting these constructs into HeLa cells, both the precursor and mature RNA expression levels were analyzed. The precursor RNA expressions from CrPV-gBglo-WT and CrPV-gBglo-P66 were similar and slight increment of precursor RNA expression from CrPV-gBglo-P66 was observed when UPF1 was down-regulated. I also found that the expression level of PTC-containing mRNAs generated from the CrPV-gBglo-P66 construct was almost the same as the expression level of normal B-globin mRNAs from the CrPV-gBglo-WT construct and PTC-containing mRNAs from CrPV-gBglo-P66 were slightly increased (Fig. 17B). These results indicate that NMD is almost completely bypassed by the CrPV-IRES element. Although the expression level of NMD bypassed PTC-containing mRNAs was as high as the expression level of their PTC free counterparts, only trace amount of mutant proteins was detected from the CrPV-gBglo-P66 construct. Moreover, down-regulation of UPF1 induced about 8-fold increase of mutant

protein expression from the CrPV-gBglo-P66 construct (Fig. 17C). These results suggest that UPF1 plays key roles as a translational repressor of PTC-containing mRNAs even after complete NMD bypass using the CrPV-IRES element.

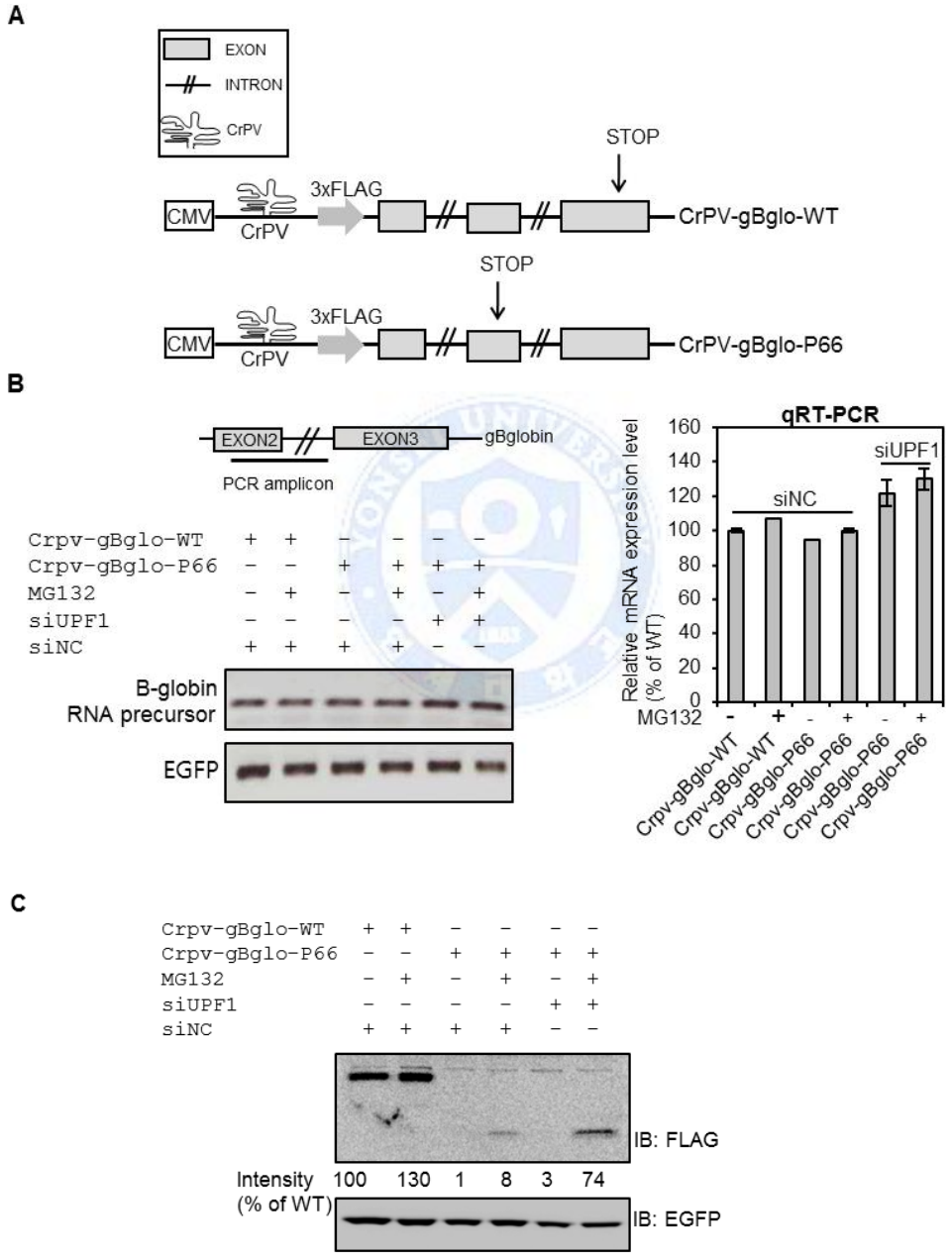


**Figure 17. UPF1 still represses translation of NMD bypassed PTC-containing mRNAs from the CrPV-IRES-gBglo-P66.** (A) The CrPV-IRES-gBglo-WT and CrPV-IRES-gBglo-P66 constructs were generated by inserting the CrPV-IRES element at the upstream site of genomic B-globin expression constructs. (B) Both the precursor RNA and mature RNA expression levels were measured by RT-PCR and qRT-PCR, respectively. Almost similar expression levels of both precursor RNAs and mature RNAs were detected from CrPV-IRES-gBglo-WT and CrPV-IRES-gBglo-P66 constructs independent of UPF1 down-regulation except for the slight increase of the mature RNA expression level from the CrPV-IRES-gBglo-P66 construct when UPF1 was down-regulated. (C) The mutant proteins were barely detected from the CrPV-IRES-gBglo-P66 construct, while the mutant mRNAs were not affected by NMD at all. More than 9-fold increases of the mutant protein expression were detected when UPF1 was down-regulated.

## IV. DISCUSSION

Quality control of mRNA expression is regulated at the RNA level by several well-known mechanisms, including NMD, NAS (nonsense-associated alternative splicing) and NMTGS (nonsense-mediated transcriptional gene silencing).[11,42,43] These mechanisms function at different times during mRNA metabolism. Of these mechanisms, NMD is a most well-documented post-transcriptional mechanism that detects PTC-containing mRNAs and subsequently inhibits translation at the initiation step and actively degrades abnormal mRNA.[44] Although NMD is an efficient mechanism for regulating abnormal PTC-containing mRNAs, there are growing evidences that suggest that NMD is not completely effective. Mutant mRNAs harboring PTCs in the last exon are not even recognized by NMD (NMD-irrelevant).[45,46] It has also been constantly reported that 10%-30% of NMD-competent PTC-containing mRNAs are resistant to NMD by unknown mechanisms and these, so called, NMD-resistant PTC-containing mRNAs exist stably.[21] More importantly, many studies have shown that NMD can be inhibited by various physiological conditions such as starvation, hypoxia, and mRNP regulations.[39,47,48] These reports ultimately suggest that there are various PTC-containing mRNAs, and

NMD are not able to remove all the PTC-containing mutant mRNAs and mutant proteins could be generated from those PTC-containing mRNAs. Consequently, it has been urged to unveil whether the mutant proteins are generated from various PTC-containing mRNAs. So far, however, specific mechanisms regarding the generation of truncated mutant proteins from PTC-containing mRNAs have not been established yet. In this study, my goal was to demonstrate whether the mutant proteins are generated from NMD-irrelevant PTC-containing mRNAs, NMD-resistant PTC-containing mRNAs and rescued PTC-containing mRNAs from NMD.

Nonsense and frameshift mutations are important inactivating mutations in the development of genetic diseases and human cancers. In particular, in human MSI-H cancers, frequent frameshift mutations have been reported in many genes.[17,19,49] When frameshift mutations induce randomized nucleotide arrangement after insertion or deletion sites, the probability of stop codon generation is approximately 3/64 (3 stop codons for every 64 codons), which means the PTCs are inevitably generated after the generation of approximately 20 abnormal neopeptides. Especially, if these mutant mRNAs contain PTCs in the last exon which NMD are not able to recognize, both mutant mRNAs and proteins are expected to be generated. Therefore, I first tried to demonstrate the generation of mutant proteins from NMD-irrelevant PTC-containing mRNAs using the MSI-H colon cancer model.

Previously, several studies have attempted to detect truncated mutant proteins derived from PTC-carrying mRNAs and clarify the roles of these mutant proteins in diseases such as cancers and genetic diseases.[14,50-55] Thus far, most of the studies have demonstrated the existence of mutant proteins at the DNA or mRNA level.[50,53] Although some studies suggested the existence of mutant proteins by *in vitro* overexpression experiments, the constructs used in the previous studies had PTCs in specific positions of the C-terminal region, which represents nonsense mutations.[54-56] In most cancers, insertion/deletion mutations usually lead to frameshift mutations

that are much more frequent and deleterious than nonsense mutations. Therefore, the identification of endogenous truncated proteins containing neopeptides from genes with frameshift mutations is more important, but this has not been studied at the protein level. I herein validated for the first time the expression of the truncated mutant proteins with neopeptides derived from NMD-irrelevant PTC-containing mRNAs from the genes with frameshift mutations.

For the validation of endogenous truncated mutant proteins from NMD-irrelevant PTC-containing mRNAs, I chose *TTK*, *TCF7L2*, and *MARCKS* which are very frequently mutated in colon cancers. The physiological roles of TTK, TCF7L2, and MARCKS have been studied in several cancers, and significant relevance of these proteins to cancer progression is well documented.[57-60] It was originally expected that significant amounts of mutant TTK might be expressed because this mutant protein has minor changes in the C-terminal region and a readthrough stop codon. Endogenous mutant TTK was barely detected by western blotting despite the relatively high mRNA expression level of *TTK* in SNUC4 cells, which have homozygous mutations (+1/−2) in *TTK*. Mutant TCF7L2 was also barely expressed, and mutant MARCKS was not expressed in the tumor cells with homozygous mutations. After proteasome inhibition by MG132 treatment, significant amounts of mutant TTK, TCF7L2 and MARCKS were detected, and I observed these dramatic increases only in the cells expressing the mutant proteins. These findings clearly indicate that the rare expression of mutant proteins of these three genes is mostly due to enhanced degradation in the proteasome pathway because I clearly demonstrated that NMD-irrelevant PTC-containing mRNAs are normally translated by polysome analysis. In addition to demonstrating the dramatic increment of mutant protein expression after proteasome inhibition, I also demonstrated the selective and heavy ubiquitination of the mutant proteins. Moreover, the colocalization of mutant MARCKS with centrosomes in the presence of MG132 suggests that mutant MARCKS proteins are actively degraded via the proteasomal machinery, which is assembled around centrosomes. By generating two different

truncated mutant MARCKS expression vectors that contained or lacked neopeptides, I demonstrated that neopeptide-containing mutant MARCKS proteins are more extensively degraded. Furthermore, I found that the neopeptide-lacking mutant proteins were relatively stable compared to the neopeptide-containing mutant proteins, and the truncated mutant displayed increased insolubility irrespective of the presence of neopeptides. All of these results explain why the mutant proteins from NMD-irrelevant PTC-containing mRNAs are barely detected, and reveal that tumor cells are protected from potentially harmful mutant proteins by the ubiquitin-mediated protein degradation mechanism when there are PTC-containing mRNAs not degraded by NMD. In terms of clinical implications of the generation and degradation of mutant proteins from NMD-irrelevant PTC-containing mRNAs, it is well known that degraded mutant proteins contribute to the formation of tumor-specific antigens and these antigens are useful targets for immunotherapy.[60] Intense peritumoral and intratumoral lymphocytic infiltration, and its association with favorable prognosis have been reported as the characteristics of MSI-H colon cancers.[17,61,62] It is also clear that the amount of intracellular mutant proteins, a substrate for tumor antigen, is closely related to the effective tumor antigen formation.[63] My results, normal protein translation from the NMD-irrelevant PTC-containing mRNAs and the generalized degradation of neopeptide-containing mutant truncated proteins, provide novel insights that the intracellular amount of mutant proteins are scant, but the degradation of neopeptide-containing mutant proteins by proteasome system is directly related to the antigen-processing and presentation by MHC class I, therefore expected to be effective tumor antigen formation. When I analyzed 19 MSI-H colon cancer tissues, I found a significant relationship between the intensity of peritumoral reaction and the mutation status of three genes [*TTK* ($p = 0.01$), *TCF7L2* ($p = 0.46$), and *MARCKS* ($p = 0.002$), Table 1]. These findings suggest that the degraded mutant proteins might be related to the regional immune responses of the tumor. A large-scale correlation study and a study on the immunostimulatory function of the neopeptides will be necessary to

determine the roles of mutant proteins in tumor antigen formation, and mutations in the other cMNR-containing genes.

After I demonstrated the generation and proteasome-mediated degradation of mutant proteins from NMD-irrelevant PTC-containing mRNAs, I further expanded my study to determine whether mutant proteins are generated from NMD-competent PTC-containing mRNAs based on recent studies that clarified the presence of NMD-resistant PTC-containing mRNAs and the fact that NMD can be inhibited by various physiological and exogenous conditions.[39] In detail, I tried to further address 1) whether mutant proteins are generated from NMD-resistant PTC-containing mRNAs and 2) whether the mutant proteins are generated from rescued PTC-containing mRNAs when NMD is inhibited. To confirm the presence of NMD-resistant PTC-containing mRNAs and their stabilities, genomic human β-globin expression vectors (gBglo-WT, gBglo-P39, gBglo-P66, gBglo-P101 and gBglo-P127) were used for this study. I observed that stabilities of *β-globin* mRNAs maintained from NMD-competent constructs (gBglo-P39 and gBglo-P66) were almost similar with the stabilities of *β-globin* mRNAs from the wild type construct (gBglo-WT) and the NMD-irrelevant constructs (gBglo-P101 and gBglo-P127) after transcription is blocked. These findings indicate that there is a stable population of NMD-competent PTC-containing mRNAs resistant to NMD. However, it is not still clear how these NMD-resistant mRNAs can be resistant to NMD. In this study, I suggested that this accumulation of NMD-resistant PTC-containing mRNAs can be explained by the capacity of NMD to remove PTC-containing mRNAs, based on the fact that the gradually increased accumulation of NMD-resistant PTC-containing mRNAs from gBglo-P66 was observed when I transfected increasing amount of gBglo-P66 plasmids into cells. When the protein expression from gBglo-P66 was analyzed, I could detect only trace amount of mutant proteins from gBglo-P66 even after MG132 treatment. Importantly, when I transfected more amount of gBglo-P66 plasmids into cells to see if the more NMD-resistant PTC-containing mRNAs are required to generate enough amount of mutant

proteins to be detected, the amount of mutant proteins was not increased even though the gradual accumulation of NMD-resistant PTC-containing mRNAs were induced through vector transfection. Subsequent polysome analysis of NMD-resistant PTC-containing mRNAs from gBglo-P66 showed that these NMD-resistant PTC-containing mRNAs from gBglo-P66 were not associated with polysomes and translationally repressed. These results are consistent with a previous report that showed most of the PTC-containing mRNAs from the NMD-competent mutant TGFBR2 construct [TGFBR2(-1)-splicing construct] are resistant to NMD and translationally repressed.[64] Even though the relative portion of NMD-resistant PTC-containing mRNAs to the entire PTC-containing mRNAs (sum of being degraded PTC-containing mRNAs and NMD-resistant PTC-containing mRNAs) are different for TGFBR2 (similar with normal *TGFBR2* mRNA expression) or β-globin constructs (about 30% of normal *β-globin* mRNA expression), both *TGFBR2* and *β-globin* NMD-resistant PTC-containing mRNAs are influenced by nonsense mediated translational repression (NMTR) which explains the translational repression of NMD-resistant PTC-containing mRNAs.

Even though NMD-resistant PTC-containing mRNAs are translationally repressed due to the NMTR effects, I could detect trace amount of mutant proteins from gBglo-P66. It has been reported that small amount of mutant proteins from the pioneer round of translation could be generated and used as potential tumor antigenic proteins.[65] Based on this idea, I tried to elucidate the involvement of the pioneer round of translation and bulky translation in the generation of mutant proteins from NMD-resistant mRNAs from gBglo-P66 and I demonstrated that the trace amount of mutant proteins from gBglo-P66 are originated from the pioneer round of translation by using 4E-BP1 and 4EGi which can specifically inhibit eIF4E dependent translation. Based on the polysome analysis results that showed rescued PTC-containing mRNAs from NMD are mostly associated with polysomes, I also tried to address if the mutant proteins could be generated from rescued PTC-containing mRNAs when NMD was inhibited. I demonstrated that both the mRNA

expression and proteins expression levels from gBglo-P66 were significantly increased when NMD was inhibited by UPF1 down-regulation. I also demonstrated that the increased protein expression level was derived from the bulky translation. By using RNAi against various core NMD or EJC factors, I further demonstrated that UPF1 plays key roles not only in NMD processes but also for the translational repression of PTC-containing mRNAs, independent of NMD. The direct involvement of UPF1 in the translational regulation of NMD bypassed PTC-containing mRNAs was also confirmed by CrPV-IRES-β-globin expression constructs. I found that PTC-containing mRNAs are still translationally repressed by UPF1 even after NMD is completely bypassed, which shows that irrespective of NMD, UPF1 works as a translational repressor for PTC-containing mRNAs. In the context of translational repression, a study reported that phosphorylated UPF1 by SMG1 interacts with eIF3, which further suppresses translation of PTC-containing mRNAs during NMD processes.[41] In this study, I extended the roles of UPF1 in terms of translational regulations of PTC-containing mutant mRNAs in the pioneer round of translation and bulky translation steps.

For the future studies, it should be confirmed whether the mutant proteins are expressed from endogenous genes with nonsense or frameshift mutations in cells and how the NMD-resistant PTC-containing mRNAs show resistance to NMD and are translationally repressed in cells. More importantly, it should be demonstrated that mutant proteins can be generated when NMD is inhibited by physiological conditions and how the physiological conditions affect UPF1 in terms of translational regulations.

In conclusion, using MSI-H colon cancer model, I demonstrated that NMD-irrelevant PTC-containing mRNAs are not degraded by NMD and these mRNAs are efficiently translated to generate truncated mutant proteins with neopeptides. However, the mutant proteins are mostly degraded by protein quality control system which is enhanced by neopeptides in the C-terminal region. On the other hand, I

demonstrated that some portion of NMD-competent PTC-containing mRNAs exist stably in cells using human β-globin constructs and these NMD-resistant PTC-containing mRNAs are translationally repressed due to NMTR effects. I also demonstrated that the trace amount of mutant proteins was generated from the pioneer round of translation step and rescued PTC-containing mRNAs by NMD inhibition were actively translated from bulky translation to generate a large amount of mutant proteins. Finally I demonstrated UPF1, a known factor involved in the translational repression, plays key roles in the translational regulation of NMD-competent PTC-containing mRNAs.

## V. CONCLUSION

To identify the fundamental molecular mechanisms for the generation of mutant proteins from NMD-irrelevant PTC-containing mRNAs, NMD-resistant PTC-containing mRNAs and rescued PTC-containing mRNAs, I analyzed the expressions of PTC-containing mRNAs and mutant proteins from endogenous mutated genes in MSI-H colon cancers and the expressions of PTC-containing mRNAs and mutant proteins from human genomic mutant B-globin expression constructs.

I found that ;

1. NMD-irrelevant mutant mRNAs containing PTC in the last exon are normally expressed in cells and translated efficiently.
2. Truncated proteins containing neopeptides are rarely detected because of extensively degradation by the ubiquitin–proteasome system which is caused by neopeptides.
3. About 30% of PTC-containing mRNAs from NMD-competent β-globin expression constructs are resistant to NMD and exist as stably as their PTC free counterparts
4. NMD-resistant PTC-containing mRNAs are translationally repressed and trace amount of mutant proteins from them is mainly generated from the pioneer round of translation step
5. Mutant proteins from the rescued PTC-containing mRNAs from NMD by UPF1 down-regulation are mostly generated from bulky translation
6. UPF1 plays key roles in the selective translational regulation of PTC-containing mRNAs depending on NMD status

In summary, I demonstrated that the differential generation and regulation of mutant proteins from NMD-irrelevant PTC-containing mRNA and NMD-competent

PTC-containing mRNAs. I demonstrated that mutant proteins from NMD-irrelevant PTC-containing mRNAs are constantly and efficiently generated but, they are mostly degraded by protein quality control system which is enhanced by neopeptides in the C-terminal region. In case of NMD-competent PTC-containing mRNAs, I demonstrated NMD-resistant PTC-containing mRNAs are translationally repressed and trace amount of mutant proteins was generated from the pioneer round of translation step. On the other hand, I found that rescued PTC-containing mRNAs from NMD are translationally active and siginificant amount of mutant proteins were generated from the bulky translation step. I also demonstrated that UPF1 plays key roles in the translational regulation of NMD-competent PTC-containing mRNAs. Taken together, my findings show the underlining molecular mechanisms of the generation of mutant proteins from the various PTC-containing mRNAs and suggest that the PTC-containing mRNAs are the useful sources for the generation of mutant proteins depending on features of PTC-containing mRNAs and NMD status (Fig. 18).

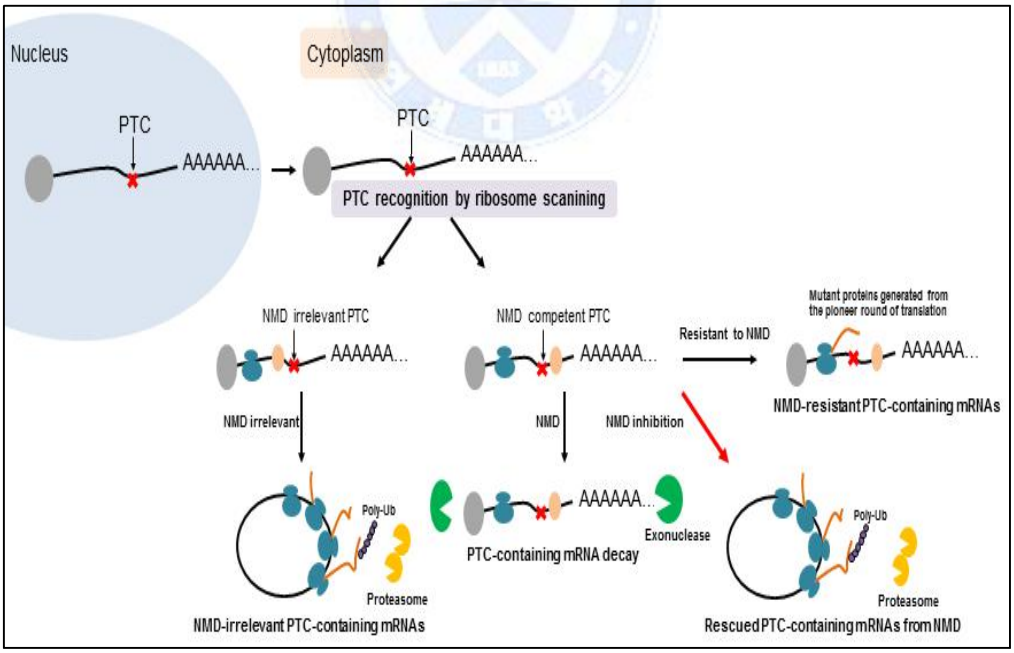



**Figure 18. Schematic models for the generation of mutant proteins from PTC-**

**containing mRNAs.** This shematic model shows the differential fates of NMD-irrelevant, NMD-resistant, and rescued PTC-containing mRNAs after their generation in terms of RNA or protein quality control mechanisms.

## REFERENCE


1. Le Hir H, Moore MJ, Maquat LE. Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. Genes Dev 2000;14:1098-108.

2. Luo MJ, Reed R. Splicing is required for rapid and efficient mRNA export in metazoans. Proc Natl Acad Sci U S A 1999;96:14937-42.

3. Nakielny S, Dreyfuss G. Nuclear export of proteins and RNAs. Curr Opin Cell Biol 1997;9:420-9.

4. Perrin-Vidoz L, Sinilnikova OM, Stoppa-Lyonnet D, Lenoir GM, Mazoyer S. The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons. Hum Mol Genet 2002;11:2805-14.

5. Wagner E, Lykke-Andersen J. mRNA surveillance: the perfect persist. J Cell Sci 2002;115:3033-8.

6. Maquat LE. When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. RNA 1995;1:453-65.

7. Maquat LE. Defects in RNA splicing and the consequence of shortened translational reading frames. Am J Hum Genet 1996;59:279-86.

8. Le Hir H, Izaurralde E, Maquat LE, Moore MJ. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J 2000;19:6860-9.

9. Metze S, Herzog VA, Ruepp MD, Mühlemann O. Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. RNA 2013;19:1432-48.

10. Kashima I, Yamashita A, Izumi N, Kataoka N, Morishita R, Hoshino S, et al. Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev 2006;20:355-67.

11. Brogna S, Wen J. Nonsense-mediated mRNA decay (NMD) mechanisms.

Nat Struct Mol Biol 2009;16:107-13.

12. Zhang J, Sun X, Qian Y, LaDuca JP, Maquat LE. At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation. Mol Cell Biol 1998;18:5272-83.

13. Zhang J, Sun X, Qian Y, Maquat LE. Intron function in the nonsense-mediated decay of beta-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm. RNA 1998;4:801-15.

14. Holbrook JA, Neu-Yilik G, Hentze MW, Kulozik AE. Nonsense-mediated decay approaches the clinic. Nat Genet 2004;36:801-8.

15. Barbier J, Dutertre M, Bittencourt D, Sanchez G, Gratadou L, de la Grange P, et al. Regulation of H-ras splice variant expression by cross talk between the p53 and nonsense-mediated mRNA decay pathways. Mol Cell Biol 2007;27:7315-33.

16. Rio Frio T, Wade NM, Ransijn A, Berson EL, Beckmann JS, Rivolta C. Premature termination codons in PRPF31 cause retinitis pigmentosa via haploinsufficiency due to nonsense-mediated mRNA decay. J Clin Invest 2008;118:1519-31.

17. Woerner SM, Kloor M, von Knebel Doeberitz M, Gebert JF. Microsatellite instability in the development of DNA mismatch repair deficient tumors. Cancer Biomark 2006;2:69-86.

18. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 1993;363:558-61.

19. Woerner SM, Yuan YP, Benner A, Korff S, von Knebel Doeberitz M, Bork P. SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. Nucleic Acids Res 2010;38:D682-9.

20. Potocnik U, Glavac D, Ravnik-Glavac M. Identification of novel genes with somatic frameshift mutations within coding mononucleotide repeats in

colorectal tumors with high microsatellite instability. Genes Chromosomes Cancer 2003;36:48-56.

21. Trcek T, Sato H, Singer RH, Maquat LE. Temporal and spatial characterization of nonsense-mediated mRNA decay. Genes Dev 2013;27:541-51.

22. Cheng J, Maquat LE. Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. Mol Cell Biol 1993;13:1892-902.

23. Belgrader P, Cheng J, Maquat LE. Evidence to implicate translation by ribosomes in the mechanism by which nonsense codons reduce the nuclear level of human triosephosphate isomerase mRNA. Proc Natl Acad Sci U S A 1993;90:482-6.

24. Lejeune F, Li X, Maquat LE. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol Cell 2003;12:675-87.

25. Pastor F, Kolonias D, Giangrande PH, Gilboa E. Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. Nature 2010;465:227-30.

26. Lengauer C, Kinzler KW, Vogelstein B. DNA methylation and genetic instability in colorectal cancer cells. Proc Natl Acad Sci U S A 1997;94:2545-50.

27. Ku JL, Yoon KA, Kim DY, Park JG. Mutations in hMSH6 alone are not sufficient to cause the microsatellite instability in colorectal cancer cell lines. Eur J Cancer 1999;35:1724-9.

28. Kim NG, Rhee H, Li LS, Kim H, Lee JS, Kim JH, et al. Identification of MARCKS, FLJ11383 and TAF1B as putative novel target genes in colorectal carcinomas with microsatellite instability. Oncogene 2002;21:5081-7.

29. Neu-Yilik G, Gehring NH, Thermann R, Frede U, Hentze MW, Kulozik AE. Splicing and 3' end formation in the definition of nonsense-mediated decay-competent human beta-globin mRNPs. EMBO J 2001;20:532-40.

30. Peixeiro I, Silva AL, Romão L. Control of human beta-globin mRNA stability and its impact on beta-thalassemia phenotype. Haematologica

2011;96:905-13.

31. Strauch KL, Beckwith J. An Escherichia coli mutation preventing degradation of abnormal periplasmic proteins. Proc Natl Acad Sci U S A 1988;85:1576-80.

32. Sieber V, Martinez CA, Arnold FH. Libraries of hybrid proteins from distantly related sequences. Nat Biotechnol 2001;19:456-60.

33. Shaid S, Brandts CH, Serve H, Dikic I. Ubiquitination and selective autophagy. Cell Death Differ 2013;20:21-30.

34. Wong E, Cuervo AM. Autophagy gone awry in neurodegenerative diseases. Nat Neurosci 2010;13:805-11.

35. Wigley WC, Fabunmi RP, Lee MG, Marino CR, Muallem S, DeMartino GN, et al. Dynamic association of proteasomal machinery with the centrosome. J Cell Biol 1999;145:481-90.

36. Bennett EJ, Bence NF, Jayakumar R, Kopito RR. Global impairment of the ubiquitin-proteasome system by nuclear or cytoplasmic protein aggregates precedes inclusion body formation. Mol Cell 2005;17:351-65.

37. Dani C, Blanchard JM, Piechaczyk M, El Sabouty S, Marty L, Jeanteur P. Extreme instability of myc mRNA in normal and transformed human cells. Proc Natl Acad Sci U S A 1984;81:7046-50.

38. Apcher S, Manoury B, Fåhraeus R. The role of mRNA translation in direct MHC class I antigen presentation. Curr Opin Immunol 2012;24:71-6.

39. Maquat LE, Tarn WY, Isken O. The pioneer round of translation: features and functions. Cell 2010;142:368-74.

40. Moerke NJ, Aktas H, Chen H, Cantel S, Reibarkh MY, Fahmy A, et al. Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. Cell 2007;128:257-67.

41. Isken O, Kim YK, Hosoda N, Mayeur GL, Hershey JW, Maquat LE. Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. Cell 2008;133:314-27.

42. Hwang J, Kim YK. When a ribosome encounters a premature termination codon. BMB Rep 2013;46:9-16.

43. Chang YF, Imam JS, Wilkinson MF. The nonsense-mediated decay RNA surveillance pathway. Annu Rev Biochem 2007;76:51-74.

44. Mühlemann O, Lykke-Andersen J. How and where are nonsense mRNAs degraded in mammalian cells? RNA Biol 2010;7:28-32.

45. Mendell JT, Dietz HC. When the message goes awry: disease-producing mutations that influence mRNA content and performance. Cell 2001;107:411-4.

46. Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci U S A 2003;100:189-92.

47. Holcik M, Sonenberg N. Translational control in stress and apoptosis. Nat Rev Mol Cell Biol 2005;6:318-27.

48. Oh N, Kim KM, Choe J, Kim YK. Pioneer round of translation mediated by nuclear cap-binding proteins CBP80/20 occurs during prolonged hypoxia. FEBS Lett 2007;581:5158-64.

49. Markowitz S, Wang J, Myeroff L, Parsons R, Sun L, Lutterbaugh J, et al. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science 1995;268:1336-8.

50. Hall GW, Thein S. Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia. Blood 1994;83:2031-7.

51. Jouanguy E, Altare F, Lamhamedi S, Revy P, Emile JF, Newport M, et al. Interferon-gamma-receptor deficiency in an infant with fatal bacille Calmette-Guerin infection. N Engl J Med 1996;335:1956-61.

52. Jouanguy E, Lamhamedi-Cherradi S, Lammas D, Dorman SE, Fondaneche MC, Dupuis S, et al. A human IFNGR1 small deletion hotspot associated with dominant susceptibility to mycobacterial infection. Nat Genet 1999;21:370-8.

53. Millar DS, Kemball-Cook G, McVey JH, Tuddenham EG, Mumford AD,

Attock GB, et al. Molecular analysis of the genotype-phenotype relationship in factor VII deficiency. Hum Genet 2000;107:327-42.

54. Fan S, Yuan R, Ma YX, Meng Q, Goldberg ID, Rosen EM. Mutant BRCA1 genes antagonize phenotype of wild-type BRCA1. Oncogene 2001;20:8215-35.

55. Sylvain V, Lafarge S, Bignon YJ. Dominant-negative activity of a Brca1 truncation mutant: effects on proliferation, tumorigenicity in vivo, and chemosensitivity in a mouse ovarian cancer cell line. Int J Oncol 2002;20:845-53.

56. Inoue K, Khajavi M, Ohyama T, Hirabayashi S, Wilson J, Reggin JD, et al. Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. Nat Genet 2004;36:361-9.

57. Chinnadurai G. CtBP, an unconventional transcriptional corepressor in development and oncogenesis. Mol Cell 2002;9:213-24.

58. Duval A, Rolland S, Tubacher E, Bui H, Thomas G, Hamelin R. The human T-cell transcription factor-4 gene: structure, extensive characterization of alternative splicings, and mutational analysis in colorectal cancer cell lines. Cancer Res 2000;60:3872-9.

59. Kalwa H, Michel T. The MARCKS protein plays a critical role in phosphatidylinositol 4,5-bisphosphate metabolism and directed cell movement in vascular endothelial cells. J Biol Chem 2011;286:2320-30.

60. Yeh YH, Huang YF, Lin TY, Shieh SY. The cell cycle checkpoint kinase CHK2 mediates DNA damage-induced stabilization of TTK/hMps1. Oncogene 2009;28:1366-78.

61. Kim H, Jen J, Vogelstein B, Hamilton SR. Clinical and pathological characteristics of sporadic colorectal carcinomas with DNA replication errors in microsatellite sequences. Am J Pathol 1994;145:148-56.

62. Kloor M, Michel S, von Knebel Doeberitz M. Immune evasion of microsatellite unstable colorectal cancers. Int J Cancer 127:1001-10.

63. Schwitalle Y, Kloor M, Eiermann S, Linnebacher M, Kienle P, Knaebel HP,

et al. Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers. Gastroenterology 2008;134:988-97.

64. You KT, Li LS, Kim NG, Kang HJ, Koh KH, Chwae YJ, et al. Selective translational repression of truncated proteins from frameshift mutation-derived mRNAs in tumors. PLoS Biol 2007;5:e109.

65. Apcher S, Daskalogianni C, Lejeune F, Manoury B, Imhoos G, Heslop L, et al. Major source of antigenic peptides for the MHC class I pathway is produced during the pioneer round of mRNA translation. Proc Natl Acad Sci U S A 2011;108:11572-7.

## ABSTRACT (IN KOREAN)

# 조기종결코돈을 포함하는 돌연변이 mRNA 의 선택적인 단백질의 생성 조절

<지도교수 김 호 근>

연세대학교 대학원 의과학과

김 원 규

일반적으로 조기종결코돈을 포함하는 돌연변이 mRNA 는 nonsense-mediated mRNA decay (NMD) 라는 기전에 의해 전사 후 조절되어 잠재적으로 위험한 돌연변이 단백질의 생성을 차단한다. 그러나 NMD 는 불완전한 기전으로 알려져 있다. 마지막 exon 에 조기종결코돈을 포함하는 돌연변이 mRNA 는 NMD 에 의해 인식되지 않아 이것들로부터 돌연변이 단백질 생성을 예측할 수 있다. 또한 약 10%~30% 정도의 NMD 에 의해 인식되는 조기종결코돈을 포함하는 돌연변이 mRNA 는 밝혀지지 않은 기전에 의해 NMD 에 저항성을 보이며 정상형 mRNA 와 유사한 안정성을 보인다. 특히 NMD 는 생리적인 또는 외부의 여러 환경에 의해 저해되는 경우가 많아, 이러한 경우 NMD 에 의해 분해되고 있던 조기종결코돈을 포함하는 돌연변이 mRNA 가 살아남을 수 있고, 이렇게 세포 내 NMD 저항성 mRNA 나 NMD 로부터 살아남은 mRNA 들은 모두 돌연변이 단백질을 만들 수 있는 잠재적 근거가 될 수 있다. 본 연구에서는 조기종결코돈을 포함하는 돌연변이 mRNA 를

돌연변이 단백질 생성 가능성과 연관하여 NMD 무관 조기종결코돈 포함 mRNA, NMD 저항성 조기종결코돈 포함 mRNA, NMD 로부터 살아남은 조기종결코돈 포함 mRNA 로 분류하고 각각의 경우 돌연변이 단백질의 생성에 대해 규명하였다. NMD 무관 조기종결코돈 포함 mRNA 로부터 돌연변이 단백질 생성에 대한 경우, 빈번한 돌연변이를 보이는 현미부수체불안정형 대장암에서 frameshift 돌연변이에 의해 마지막 exon 에 조기종결코돈을 갖는 29 개의 유전자를 발굴하고 이들 중 3 개의 유전자 TTK, TCF7L2, MARCKS 를 선정하여 각각 돌연변이 mRNA 와 단백질의 발현을 분석하였다. 돌연변이를 TTK, TCF7L2, MARCKS 유전자로부터 발현되는 NMD 무관 조기종결코돈 포함 mRNA 는 모두 NMD 에 의해 분해되지 않고 안정적으로 존재하였으나, 돌연변이 단백질은 거의 확인되지 않았다. Genomic MARCKS 발현 벡터를 이용한 Polysome 분석을 통해 NMD 무관 조기종결코돈 포함 mRNA 가 정상적으로 번역되고 있으며, 이러한 돌연변이 단백질은 생성 후 proteasome 시스템에 의해 빠르게 분해되고 있다는 사실을 규명하였다. 또한 이러한 돌연변이 단백질의 분해 촉진은 c-terminal 부위에 존재하는 비정상형 neopeptide 때문이라는 사실을 확인하였다. 한편, NMD 저항 조기종결코돈 포함 mRNA 와 NMD 로부터 살아남은 조기종결코돈 포함 mRNA 로부터의 돌연변이 단백질 생성을 규명하기 위해 조기종결코돈을 포함하는 사람 β-globin 유전자 발현 벡터를 이용하여 mRNA 와 단백질의 발현을 분석하였다. 그 결과, 약 30% 정도의 NMD 저항 조기종결코돈을 포함하는 mRNA 가 NMD 에 저항성을 보이며 정상형 mRNA 만큼 안정적으로 존재하는 것을 확인하였다. 그러나 이렇게 안정한 NMD 저항 mRNA 로부터는 오직 극미량의 돌연변이 단백질만을 확인할 수 있었고, 이는 NMD 저항 mRNA 의 번역이 억제되어 있으며, 극미량의 돌연변이 단백질은 pioneer round of translation 중 scan 과정에서 생성 되었다는

사실을 규명하였다. 또한 NMD 의 주요 구성 요소인 UPF1 을 저해하여 NMD 를 억제하였을 때, 조기종결코돈을 포함하는 mRNA 가 살아남는 것을 확인하였고, bulky translation 과정으로 들어가 많은 양의 돌연변이 단백질을 생성한다는 사실을 규명하였다. 또한 이러한 조기종결코돈을 포함하는 돌연변이 mRNA 의 번역을 억제하는 전반에는 UPF1 이 핵심적인 역할을 한다는 사실을 확인하였다. 이러한 결과들은 조기종결코돈을 포함하는 mRNA 들이 NMD 에 관련 또는 무관하게 돌연변이 단백질을 만들어낼 수 있는 잠재적 물질이라는 사실을 보여준다.

핵심되는 말: 절단형 돌연변이 단백질, 비정상 펩타이드, frameshift 돌연변이, nonsense 돌연변이, 조기종결코돈, nonsense-mediated mRNA decay (NMD), NMD-저항, NMD-무관, 구조된 PTC-containing mRNAs

## PUBLICATION LIST


1. Song M, Kim H, Kim WK, Hong SP, Lee C. High expression of AT-rich interactive domain 3A (ARID3A) is associated with good prognosis in colorectal carcinoma. Ann Surg Oncol 2014;21 Suppl 4:S481-9.

2. Park M, Kim M, Hwang D, Kim WK, Kim SK, Shin J, et al. Characterization of gene expression and activated signaling pathways in solid-pseudopapillary neoplasm of pancreas. Mod Pathol 2014;27:580-93.

3. Park M, **Kim WK**, Song M, Kim H, Nam HJ, Baek SH. Protein kinase C-δ-mediated recycling of active KIT in colon cancer. Clin Cancer Res 2013;19:4961-71.

4. **Kim WK**, Park M, Kim YJ, Shin N, Kim HK, You KT, et al. Identification and selective degradation of neopeptide-containing truncated mutant proteins in the tumors with high microsatellite instability. Clin Cancer Res 2013;19:3369-82.

5. **Kim WK**, Yang HK, Kim H. MicroRNA involvement in gastrointestinal stromal tumor tumorigenesis. Curr Pharm Des 2013;19:1227-35.

6. Bertucci F, Finetti P, Ostrowski J, Kim WK, Kim H, Pantaleo MA, et al. Genomic Grade Index predicts postoperative clinical outcome of GIST. Br J Cancer 2012;107:1433-41.

7. **Kim WK**, Park M, Kim YK, Tae YK, Yang HK, Lee JM, et al. MicroRNA-494 downregulates KIT and inhibits gastrointestinal stromal tumor cell proliferation. Clin Cancer Res 2011;17:7584-94.

8. Shin N, You KT, Lee H, Kim WK, Song M, Choi HJ, et al. Identification of frequently mutated genes with relevance to nonsense mediated mRNA decay in the high microsatellite instability cancers. Int J Cancer 2011;128:2872-80.